Essays in Development Economics

Laura Montenbruck

Acknowledgments

Contents

Preface

This dissertation uses empirical, mostly experimental, methods to study questions in development economics, with a focus on behavioral responses to the provision of information. The thesis consists of three self-contained chapters.

Chapter 1 is titled "Fiscal Exchange and Tax Compliance: Strengthening the Social Contract Under Low State Capacity". In this first chapter, I conduct a field experiment randomizing the provision of specific, targeted information about public service delivery across 5,494 property owners and tenants in Freetown, Sierra Leone. I study whether receiving such information affects property tax compliance. I depart from existing work investigating the effects of very general information about the link between taxation and public service delivery by providing individuals with specific examples of services that have been delivered by their local government. In addition, individuals are informed about the location at which these services have been provided. This makes public service provision both tangible and verifiable, thus increasing the credibility of the government's effort to fulfill its part of the social contract.

I find that receiving specific information increases property tax payments by 20% on average. The effect is driven by increases in tax compliance on both the extensive and intensive margin. Among poorer individuals, the likelihood of paying taxes increases by 7-16 percentage points in response to information about services that are both geographically accessible and reflect the citizens' most urgent needs, suggesting a benefit-based approach to taxation. These effects are substantial compared to an average compliance rate of around 15% for this subset of individuals. Revenue effects are largely driven by richer individuals, who depend less on the provision of public services, and for whom the treatment seems to act as a more general signal of government capacity. These results show that a simple information intervention has the potential to substantially increase tax payments, and as such government revenue. They also suggest that information about public service provision affects tax compliance through different channels depending on the extent to which individuals rely on the latter.

Chapter 2, "The Power of Faith: Effects of an Imam-led Information Campaign on Labor Supply and Social Interactions", is joint work with Alexandra Avdeenko, Jakob Gärtner, Marc Gillaizeau, Ghida Karbala, Giulia Montresor, Atika Pasha, and Galina Zudenkova. In this

chapter, we investigate whether information spread by religious leaders can effectively change individuals' attitudes towards health risks and their compliance with recommended health measures. We conduct a randomized controlled trial with more than 6,000 NGO-beneficiaries in rural Pakistan, randomizing a remote awareness campaign and its endorsement by religious leaders in the context of the COVID-19 pandemic. Treated individuals either receive information via a phone call from employees of the NGO, or via a combination of phone call and loudspeaker announcements made at a local mosque. We study the effects of the two treatment arms on beliefs and preventive behavior related to the virus.

We find that the phone campaign endorsed by religious leaders via mosque loudspeakers leads to a significant decrease in labor supply. The effect is driven by male respondents, whose likelihood of working outside home decreases by 4.8 percentage points in response to the treatment. We also observe a significant reduction in social interactions among male individuals who received the combined intervention. Our results on knowledge and perceptions related to COVID-19 suggest that these effects are driven by an increase in the perceived risk of transmitting the virus to others, indicating an altruistic motivation behind the more cautious dealing with the pandemic. The effects of the phone and loudspeaker campaign differ significantly from those of the pure phone call treatment, which does not affect behavior at all. These results suggest that the authority and credibility of religious leaders played a crucial role in motivating pro-social, health-conscious behavior.

Chapter 3 is titled "The Economics of Women's Rights" and is co-authored with Michèle Tertilt, Matthias Doepke, and Anne Hannusch. In this chapter, we investigate the drivers behind reforms leading to increases in women's rights across time and space. To do so, we distinguish between four types of rights: economic, political, labor, and body rights. We document how these rights have evolved over the past 50 years across countries and present a theoretical framework that highlights four important economic channels for political-economy models, which can explain support for and opposition to women's rights. We show that economic factors play a key role in the evolution of women's rights. Importantly, variables proxying parental altruism—i.e., fathers' support for women's rights as a reflection of care for their children's well-being—and an income channel—i.e., support for women's rights based on the expectation that they would positively affect a household's total resources—are highly correlated with all aspects of women's rights.

Chapter 1

Fiscal Exchange and Tax Compliance: Strengthening the Social Contract Under Low State Capacity

Abstract: This article provides evidence that increased salience of public service provision can strengthen the social contract and increase tax compliance in a low-capacity setting. I conduct a field experiment randomizing information about public service provision across 5,494 property owners and tenants in Freetown, Sierra Leone. Receiving information increases property tax payments by 20% on average. The effect is driven by increases in tax compliance on both the extensive and intensive margin. Residents of low-value properties are 7–16 percentage points more likely to pay taxes when informed about public services that are both geographically accessible and respond to the citizens' most urgent needs, suggesting a benefit-based approach to taxation. Revenue effects are largely driven by residents of high-value properties, who depend less on the public provision of services, and for whom the treatment seems to act as a more general signal of government performance.

1.1   Introduction

The ability to tax is central to the process of state development (Besley and Persson, 2013). Tax revenue enables the provision of public services and allows countries to form capital from domestic sources (Kaldor, 1963). Despite its crucial role, many low-income countries fail to generate adequate levels of tax income.[1] The effectiveness of states—including their ability to tax—is associated with two key factors: coercive power (e.g., Hobbes, 1651), and the existence of a social contract (e.g., Locke, 1690; Rousseau, 1762). Low administrative capacity limits the ability to enforce tax policies in low-income countries (Besley and Persson, 2013). A salient contractual relationship between citizens and the state that acknowledges and fulfills reciprocal obligations, however, has the potential to encourage "quasi-voluntary" tax compliance (Levi, 1988).

In her historical account of the power to tax, Levi (1988) argues that citizens' quasi-voluntary tax compliance—the decision to pay taxes in a setting where non-compliance is technically punishable by law—depends crucially on their confidence that the state fulfills its part of the tax contract. Embedded in a growing literature on tax morale, recent empirical work has investigated whether highlighting the link between taxation and public service provision affects tax compliance (see, e.g., Luttmer and Singhal, 2014). Focusing mainly on high- and middle-income countries, these studies either emphasize the link between taxation and public service provision theoretically, or provide aggregate information about the use of tax revenue.[2] However, in low-capacity settings, where trust in the reciprocity of the government tends to be low (Prichard, 2017; Dom et al., 2022), and it is less recognized that taxes finance important public services (Mascagni, 2018), such general information may not be enough to credibly signal that the government fulfills the social fiscal contract.

In this paper, I test whether specific, targeted information about recently provided public services increases property tax compliance in a low-capacity setting. The study is situated in Freetown, Sierra Leone, where property tax is a major source of municipality level revenue.[3] The Freetown City Council (FCC) is making substantial investments to improve public service provision throughout the city and property owners report being willing to pay taxes in exchange for the provision of public services (Grieco et al., 2019). However, despite being registered,

---

[1]   Whereas the tax-to-GDP ratio amounts to 33.5% in OECD countries, it is as low as 16% on average among 31 African countries for which tax data is available. With 11%, the ratio is even lower in Sierra Leone (OECD, 2022).

[2]   "Your tax payment contributes to the funding of publicly financed services in education, health and other important sectors of society." (Bott et al., 2020) is an example of a theoretical message emphasizing fiscal exchange. "In the first 6 months of this year, CVP's collection contributed to placing 28 new streetlights, water connections in 29 streets and sewerage networks in 21 blocks." (Castro and Scartascini, 2015) is an example of aggregate information about the use of tax revenue.

[3]   In 2021 and 2022, property tax accounted for 46% and 56% of municipality level revenue, respectively. These numbers only account for realized property tax payments.

only few property owners pay taxes.[4]

Against this background, I design an intervention that randomly provides residents of registered properties with specific and geographically targeted examples of public services delivered by the local government. To do so, I verify and geo-reference all locations at which public services have been provided in the three years prior to the intervention. I create a database including a total of 277 services aiming, among other things, at improving the quality of education, access to water, and sanitation. I link this information to administrative data available from the property tax registry, including the specific location of each property. This allows me to inform individuals about the closest location at which a certain type of service has recently been delivered.[5] In particular, treated individuals are informed about (a) a specific type of service, (b) an example location at which the service has been provided, and (c) the approximate time it takes them to get to said location from their house.[6]

To broaden our understanding of the mechanisms at play, I elicit and study the role of public service preferences. Laboratory experiments suggest that providing services in line with taxpayers' elicited preferences has the potential to increase tax compliance (Alm et al., 1993; Lamberton et al., 2018), supporting a theory of benefit-based tax compliance. However, evidence on the real-world link between public service preferences and tax payments is limited.[7] Thus, during a baseline interview, I ask individuals about both their most and least preferred type of service to be provided by the local government. More than half of all individuals indicate improved access to water as their top priority.[8] Notably, this priority coincides with the municipality's current focus on providing access to water.

I use elicited preferences to design two different treatment arms: In the first group, individuals receive information about the type of service they considered their most preferred service. I refer to this treatment as the *match* treatment. In the second group, I inform individuals about a service of their least preferred type. I refer to this treatment as the *mismatch* treatment, as it creates the largest mismatch possible between the type of service individuals receive information about and the type of service they consider their top priority. I compare these two treatment

---

[4] In 2021, the year prior to the intervention, property tax was paid for only 22% of all registered properties.

[5] To account for the fact that these locations are not selected randomly—they may be selected because of a previous lack of services, or reinforce existing differences in service provision across the city—and in the absence of detailed data on public services beyond those recently provided, I account for geographical differences in the base level of public services using ward fixed effects throughout my analysis.

[6] While the example location is the closest one at which the respective type of service has been provided, individuals are not explicitly informed about this fact.

[7] Survey experiments show that eliciting preferences can increase tax morale (Sjoberg et al., 2019; Casal et al., 2016). Khan et al. (2022) tests the effect of providing services aligned with individuals' preferences on property tax payments in Pakistan with no final results available yet. By contrast, I provide *information* about the provision of public services aligned or misaligned with preferences.

[8] As compared to this, slightly less than 50% of individuals claim disaster management and prevention or environmental management to be their lowest priority.

groups against a pure control group of individuals who receive no information at all.[9]

Based on the above-mentioned evidence, I expect that receiving information about services aligned with one's preferences has a positive effect on tax compliance. However, it is ex ante less obvious what effect information about one's least preferred service will have. If individuals have zero preference for the provision of their least preferred service, informing individuals about the latter will likely not affect tax compliance. If, however, individuals have a positive preference for the provision of their least preferred service, the effect can be expected to be positive and increasing in the closeness of preferences. The effect of the mismatch treatment may also interact with prior knowledge about the provided service, and as such the extent to which the treatment provides new information. For instance, individuals may choose not to acquire (costly) information about a service that they have a low preference for and therefore may not be aware of its provision. In this case, the mismatch message would on average provide more new information, thus shifting priors and potentially increasing tax compliance. Whether the mismatch treatment does or does not affect tax compliance, and if so, through what channel, will be investigated empirically.

I implement this intervention in a large-scale randomized controlled trial (RCT) with 5,494 residents of registered properties. Properties are sampled from the universe of registered properties, with eligibility being restricted to properties owned by a single-property owner for which at least one phone number is available (N=46,844).[10] I randomly allocate one third of these individuals to the control group, one third to the *match* and one third to the *mismatch* group. The treatment message is conveyed during a short phone call conducted by a qualified enumerator. Right after the call, treated individuals receive a short message (SMS) summarizing the main information provided during the call.

To investigate the effects of this intervention on property tax compliance, I use administrative tax payment data made available by the Freetown City Council (FCC). This data allows me to capture actual payments made rather than declarations of income or revenue, which determine the tax base, but not actual compliance. I complement this analysis with information from an endline survey with a subset of individuals. This data allows me to look at intermediary outcomes related to the reciprocity channel, as well as to investigate potential alternative mechanisms through which the treatment could raise compliance, such as increases in perceived government capacity to track and punish evaders or awareness of one's tax duty.

---

[9] While this does not allow me to directly compare the effects of the fiscal exchange message against that of alternative messages such as reminders or deterrence messages, I look at these channels as potential drivers of the observed effects using suggestive evidence from an endline survey as well as results on the heterogeneity of effects.

[10] Single-property ownership is approximated using the names and phone numbers associated with each property. A more detailed description of the process can be found in Section 1.4. These properties represent approximately 45% of all registered properties.

I find that specific and geographically targeted information about recently provided public services significantly increases tax compliance by the payment deadline. Approximately six months after the intervention, treated individuals pay on average USD 1.5 (SLE 23.8) more taxes, a 20% increase as compared to the control group mean.[11] For the 3,660 treated individuals, the treatment increased timely payments by more than SLE 87,000, or approximately USD 5,600. Extending the intervention to all 46,844 eligible properties could increase tax revenue by more than USD 71,700. With a conservative per-individual cost of SLE 14.2 (including the cost of mapping all services as well as enumerator and phone costs for implementing the intervention), the intervention generates 1.7 times its cost. Almost half of the cost is driven by expenses for phone calls and SMS—which are particularly expensive in Sierra Leone. As the government often reaches more advantageous agreements with network providers, this is likely a lower bound of the return to providing public service information at scale.[12] These findings suggest that specific information about the provision of public services can be an effective tool to increase tax compliance.

To get a more nuanced understanding of the mechanisms driving these effects, I look at two different types of individuals: poor individuals, who reside in low-value properties, and rich individuals, who reside in high-value properties. Poor individuals rely heavily on public service provision and have lower access to public services.[13] Accordingly, these individuals react to the treatment message in different ways as compared to rich individuals. Residents of low-value properties increase tax payments in response to information aligning with their most preferred type of service, indicating a benefit-based approach to tax compliance as expected. The effects seem to be driven predominantly by increases on the extensive margin of tax compliance. Residents of high-value properties, on the other hand, increase tax payments in response to the mismatch treatment. As these individuals rely on average less on public service provision, tax compliance seems to be less closely related to personal benefit from the services they are informed about. Indicative evidence suggests that these individuals consider the treatment as a positive signal of government performance more generally. Whereas the mismatch treatment shifts individuals' priors about the government's performance in providing public services, the match treatment does not. This suggests that individuals may have been more informed about their most preferred service ex ante and explains why the match treatment does not affect tax compliance among residents of high-value properties.

The differential effects of the two treatment arms provide important insights into the role

---

[11] Nominal exchange rates for September 30, 2022 taken from https://www.exchangerates.org.uk/USD-SLL-spot-exchange-rates-history-2022.html.

[12] In previous interventions conducted by the government itself, network providers even allowed free usage of their network.

[13] Rich individuals, on the other hand, not only have higher access, but also often provide services privately and therefore depend less on the government's fulfillment of the social contract.

of personal benefit from the provision of public services for tax compliance. Beyond stated preferences, however, several other factors can determine the extent to which individuals benefit from the services they are informed about. Two aspects are of particular interest given the design of the intervention: *access* and *contextual importance* of specific types of services. Both factors implicitly form part of the treatment message: information about access can be derived from the measure of *distance* (approximated by commuting time) stated in the message; considered independently from preferences, different *types of services* are likely perceived as differentially beneficial among all individuals facing the same environment. In the context of Freetown, I focus on the role of access to water—a service that can unambiguously be considered essential—as compared to any other type of service. Not only is water a basic need and prerequisite for being able to benefit from other services. It is also a service that is poorly provided in the city of Freetown.[14]

I exploit the variation in distance to and type of service across individuals and within treatment groups to investigate heterogeneity in treatment effects with respect to the individualized content of the message. By design, the composition of the types of services individuals are informed about, as well as distance to that service, vary across treatment arms. However, both treatment arms as well as the control group are balanced in terms of distance to the closest service of each type and well-balanced with respect to the share of individuals indicating a specific service as their most or least preferred type.

Disentangling the effects of the treatment depending on the message's content reveals that the effects on tax compliance are driven by individuals who receive information about the provision of water close to where they live. In line with the idea that personal benefit of public service provision matters predominantly for those who depend on public provision of services, the specific content of the treatment message correlates with tax compliance only among residents of low-value properties. For this subgroup of individuals, the likelihood of making a payment increases by 7-16 percentage points, or 62–134% as compared to the respective control group, when informed about the provision of water at an accessible distance.[15] These results underline the importance of considering multiple layers of perceived benefit of public service provision, including not only stated preferences, but also access and the more general valuation of certain types of services.

To demonstrate that treated individuals indeed increase tax compliance based on a fiscal exchange (reciprocity) motive, I provide indicative evidence ruling out two alternative mecha-

---

[14] Indicative evidence suggests that around 39% of registered properties in Freetown have no access to water at all, and only around 45% are connected to the piped water system (Grieco et al., 2019). The situation is even more dire when considering the entire population of Freetown, with only 5% of individuals having access to piped water (Freetown City Council, 2019).

[15] As the intensive margin effects are driven almost exclusively by high-value properties, the content of the message does not correlate with these effects.

nisms: increases in perceived extractive capacity of the government, and increases in tax aware-
ness. To show the former, I look at self-reported outcomes related to the perceived capacity of
the government to collect taxes and punish evaders. I find no evidence that increased fear of
being caught or punished when evading drives the effects on tax compliance. Importantly, this
is also true for residents of high-value properties who received the mismatch treatment, sug-
gesting that the effect of the mismatch treatment cannot be explained by fear of enforcement.
The effects do not seem to be driven by increased awareness of one's tax duty either. Intuitively,
this is because the treatment message was designed to focus on public service provision without
mentioning tax obligations.[16]

Finally, I show that the effects on tax compliance persist over the entire tax cycle, suggesting
that treated individuals do not only pay taxes earlier, but actually pay more taxes. By the end of
the 2022 tax cycle, individuals living in low-value properties who received information about
water at close distance are significantly more likely to have paid any tax. This translates into a
21% increase in the average amount of tax paid among individuals in the match group—who
by design are about 6 times as likely to receive information about water as individuals in the
mismatch group. Among residents of high-value properties, those who received the mismatch
treatment pay on average 29% more taxes during the entire 2022 tax cycle than those in the
control group.

This paper contributes to several strands of literature. First, I add to a growing literature
on property taxation in low-income countries. As urbanization is expected to increase dras-
tically, cities need to generate more revenue to provide adequate levels of public services.[17]
The taxation of properties has been postulated to be the "best available instrument for funding
local governments, because they are generally progressive, economically efficient, and poten-
tially closely linked to local services" (Moore et al., 2018, p.151). However, countries face
multiple challenges along the process, including the registration and valuation properties, deter-
mining optimal tax rates, and incentivizing compliance. Recent work has therefore investigated
how countries can successfully register taxpayers (Weigel, 2020), assess the value of proper-
ties (Knebelmann and Pouliquen, 2023), determine the optimal tax rate (Bergeron et al., 2023;
Brockmeyer et al., 2023), or increase property tax compliance (Khan et al., 2019; Krause, 2020;
Balán et al., 2022; Collin et al., 2022; Fernandez et al., 2023). I contribute to this literature by
providing evidence on the effectiveness of specific, targeted information about public service

---

[16] This conclusion is also supported by several other pieces of evidence. First, I show that the treatment effect is
driven by individuals who previously contributed to the provision of community services, i.e., individuals who
care disproportionally about the provision of public services. Second, the fact that the message's specific con-
tent matters contradicts the hypothesis that being treated increases tax compliance via increased tax awareness
independently of the fiscal exchange mechanism.

[17] In 2018, approximately 42.5% of the African population lived in urban areas. This number is expected to
increase to 58.9% by 2050 (United Nations, 2019).

provision on tax compliance in a low-enforcement setting.

Second, I add to the literature on the effects of messages emphasizing the link between tax-ation and public service provision, thereby trying to strengthen the social fiscal contract and nudge tax compliance.[18] Previous studies from high- and middle-income countries have found that these types of messages are generally insufficient to increase compliance.[19] Evidence from low-income countries is more scarce, but largely supports these findings. In particular, stud-ies from Rwanda and Uganda find no effects of such general messages on average (Mascagni and Nell, 2022; Cohen, 2023).[20] Importantly, previously studied messages are general in na-ture and merely inform individuals that tax revenue is necessary to finance public services or provide aggregate, hard-to-verify information about the use of tax revenue. By contrast, I pro-vide individuals with targeted information about services delivered by their local government. This distinction is substantial as citizens often lack awareness of the public services provided to them, or have biased beliefs about how the state spends its revenue (Khan et al., 2022; Giac-cobasso et al., 2023) and confidence in the reciprocity in the government is low (Prichard, 2017; Moore et al., 2018; Dom et al., 2022). This can explain why more generic information may not be enough to motivate them to pay taxes. Crucially, by providing individuals with specific ex-amples of services delivered by their government, my treatment makes the government's effort in fulfilling its part of the social fiscal contract more salient. My work shows that providing specific information about public services delivered by the local government increases property tax compliance. This finding adds important context to the more pessimistic previous results on the effects of messages highlighting the link between tax compliance and public service provi-sion in a more abstract way. These insights are relevant for policymakers who weigh the costs and benefits of different policies aimed at raising tax compliance.

Finally, I add to our understanding of the drivers of perceived fiscal exchange by disen-tangling the different components of the treatment message. Evidence from lab (Lamberton et al., 2018) and survey experiments (Sjoberg et al., 2019; Abbiati et al., 2020) suggests that individuals are more willing to pay taxes when asked about their preferences on how to use tax revenue. Similarly, lab (Alm et al., 1993; Casal et al., 2016) experiments show that allowing

---

[18] The role of reciprocity has also been studied in theoretical work. For instance, Besley (2020) proposes a model linking tax compliance to reciprocal motivations.

[19] E.g., Blumenthal et al. (2001); Castro and Scartascini (2015); De Neve et al. (2021) and Bergolo et al. (2023) find no effects of messages emphasizing fiscal exchange. Hallsworth et al. (2017) and Bott et al. (2020) find positive effects on the timing of payments and income declared in the UK and Norway, the latter pointing out favorable conditions of high trust in the government and high initial tax compliance as potential drivers of the effects. Hernandez et al. (2017) even finds negative effects on tax compliance in Poland. Antinyan and Asatryan (2020) conducts a meta-analysis of more than 40 interventions and concludes that non-deterrence tax messages, including fiscal exchange messages, are on average ineffective.

[20] One exception is Collin et al. (2022), which finds positive and economically significant effects of reciprocity messages in Tanzania. This may indicate a more favorable environment for fiscal exchange messages, or could be related to the fact that around 80% of treated had not received a tax bill prior to the data collection.

individuals to influence the allocation of revenue to services they prefer over others increases tax contributions. Yet, evidence from the field confirming these findings is scarce (Khan et al., 2022; Giaccobasso et al., 2023). I show that individuals who depend substantially on public provision of services indeed react to information on services aligned with their elicited preferences. However, these effects are largely driven by information on services which align with the more general needs of the population (water) as well as by accessibility of these services. While the former suggests that the elicitation of individual preferences may be of secondary order as long as the government can proxy citizen needs on an aggregate level, the latter shows that characteristics of public service provision which have so far not been incorporated in lab experiments are crucial indicators of the willingness to pay taxes in real life settings.

1.2 Institutional Background

Property tax reform   In 2019, the city of Freetown introduced a new property tax system. Prior to this reform, the valuation of properties was done manually and was purely based on the floor area of a property. Whereas this method was simple, it failed to account for important property characteristics such as the location and quality of a building, and as such was both inaccurate and regressive (Jibao, 2017). The new 'points-based method' combines information on the surface area with additional observable characteristics of the property and surrounding area, assigning a standard number of points for the former and adding or deducting points for positive or negative property features (Fish, 2018). Using a model informed by the characteristics and rental value of a subset of properties, this information is then translated into an estimated property value (Grieco et al., 2019). Importantly, the assessed value of a property also depends on access to public services, such as water, drainages, and streets. Thus, individuals with better access to public services have on average higher tax liabilities.

The development of the new property assessment scheme went hand in hand with the registration of all qualifying buildings into the property registry.[21] Buildings were identified using satellite imagery and were then visited by enumerators who geo-referenced them and collected the data on observable characteristics necessary to estimate the property value. During this process, more than 50,000 new properties were registered, increasing the tax base to roughly 110,000 taxable properties throughout the city (Kamara et al., 2020).

The introduction of the new property tax system had considerable effects on the tax liabilities of property owners. Whereas the tax rate for properties in the first two quintiles of average tax payable decreased by more than half, tax liability more than tripled for the top 20% of prop-

---

[21] This does not include non-permanent and uncompleted properties and illegally constructed properties. Furthermore, public buildings such as clinics, schools, universities, and places used for religious worship are exempted from the property tax.

Table 1.1: Change in tax liabilities by property value quintiles (in USD)

Average tax payable	Existing system	New system	Average change
1st Quintile	$14.33	$4.31	-70%
2nd Quintile	$15.85	$9.48	-40%
3rd Quintile	$16.10	$17.40	+8%
4th Quintile	$23.38	$36.94	+58%
5th Quintile	$41.64	$142.25	+242%

Source: Table adapted from Kamara et al. (2020)

erties (see Table 1.1). However, despite increased transparency and accuracy of the property assessment, as well as the introduction of penalties for late payers, tax compliance remains low. In 2021, i.e., after the reform, but prior to the intervention examined in this paper, only 22% of property owners paid property taxes. At the same time, property tax is the most important source of income for the municipality. 46% of local government revenue consisted of property taxes in 2021, increasing to 56% in 2022 (see Appendix Figure A1). Calculating potential tax revenue—that is, tax revenue if all annual tax liabilities were paid—property tax would make up around 80% of total revenue.[22] This underlines the importance of property tax and points to the need of testing additional means to increase tax compliance, such as messages related to tax morale.

Payment logistics   Property tax payments for a specific tax cycle are made in the same tax year, i.e., property tax payments for the 2022 tax cycle are made in 2022. Between January and March, rate demand notes (RDNs) are issued and delivered to the properties via mail (printed, not online). RDNs specify the current tax liability including arrears, penalties, and credit. They also contain information about the value assessment of the property. Tax payments can be made either directly at the Freetown City Council or at a number of bank branches throughout the city. Payments can be made in installments of any size, which will be credited to the outstanding tax liability.

Individuals who pay before March 31 receive a 5% credit on the amount paid, which is deducted from the tax due in the following year. The deadline for tax payments is September 30. Individuals who exceed this deadline have to pay a penalty of 5%, which is added to the tax payment due in the following year. Despite the theoretical penalty mechanism, enforcement of payments is extremely low. As of now, liabilities and penalties are mostly accumulated over time without actual consequences for the debtor. This is somewhat different for properties with very high value, for which, anecdotally, additional efforts to collect taxes, especially after the

---

[22] This includes all tax liabilities from domestic, commercial, and institutional properties and holds all other revenue equal.

payment deadline, are higher. By law, properties can be seized if individuals are non-compliant despite official warnings and penalties. Again, this is not something that regularly happens in practice.

1.3   Data and Descriptives

I use both administrative records and survey data. I describe this data briefly here, and in more detail in Appendix 1.A. Through an agreement with the Mayor of Freetown and the Freetown City Council (FCC), I had access to the database of registered properties in Freetown. This database was used to sample study participants and track property tax payments over time. To inform the content of the treatment message, I generated a dataset of recently provided public services including geo-referenced locations. Additionally, baseline data was collected from all study participants to elicit preferences for local government-provided services and to ascertain beliefs related to taxation and governance. For a subset of individuals, I collected endline data to explore potential alternative drivers of tax compliance.

The study was carried out between February and April 2022 and was conducted purely over the phone. Figure 1.1 illustrates the timeline of the data collections as well as the intervention. The baseline survey was conducted between February 11 and March 4, information was provided to treated individuals between March 29 and April 12, and endline interviews started on April 6 and ended on April 16.[23]

Figure 1.1: Timeline of data collection (2022)
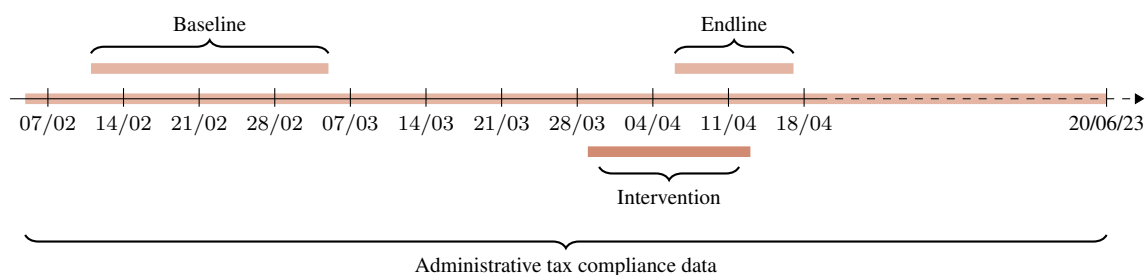
Administrative data   The database of registered properties contains all properties in Freetown which are legally obliged to pay property tax, including domestic, commercial, and institutional

---

[23] Whereas the timing of the intervention and the endline data collection overlap overall, there are on average 9 days between the intervention and the endline survey for treated individuals, and 96% of endline interviews were conducted between 8 and 12 days after the treatment took place. See Appendix Figure A2 for the exact distribution of days between intervention and endline survey in the treatment group. Note that the endline data collection was organized in a way that aimed at collecting equal shares of control group and treatment group data every day to ensure comparability.

buildings. In January 2022, the registry counted 103,407 properties. Besides information about the value of properties and the associated property tax rate, the database contains information about the exact location of the property (address and geo-referenced location) as well as contact information (phone numbers) of the owner or another individual associated with the property, e.g., a tenant, a family member of the owner, or a caretaker.[24] The registry records every tax payment made over the course of the year together with the date on which the payment was made. I have access to this data for the 2021 and 2022 tax cycles as well as part of the 2023 tax cycle, including payments until June 20, 2023.

I restrict the set of properties to properties used for domestic purposes, which are owned by a single-property owner, and for which at least one phone number is available. Multiple-properties ownership is used as a proxy for wealthier households. Since these households generally rely less on *public* service provision and instead tend to provide services *privately*, they are expected to be affected less by information about publicly provided services. To concentrate efforts on the population which I anticipate to respond most to the treatment, these properties are excluded. This results in a set of 46,844 eligible properties, out of which 15,000 were randomly drawn to build the sample frame.[25] The sample frame exceeded the planned amount of 7,000 interviews to account for potential non-response.

Geo-mapping of public services   To provide accurate and targeted information about recent public service provision, I used a publicly available information brochure issued by the Freetown City Council, which lists public services provided between 2019 and 2021 for each ward of Freetown.[26] Using this brochure, enumerators were sent to locate the mentioned services to (a) confirm their existence and (b) save their exact geo-referenced location.[27] In total, 277 services were geo-referenced. This list includes all services provided and listed by the FCC for which existence could be confirmed by the enumerators.

Based on the services provided by the local government, I identified eight key areas of municipality level public service action: (1) access to water, (2) disaster management and pre-

---

[24] The desktop version of the property registry also contains detailed information about the characteristics of each property used to assess the property value. Unfortunately, I do not have access to this data in an aggregated manner.

[25] Note that in few cases, several properties use the same phone number without necessarily being owned by the same person. To avoid sampling the same phone number twice, properties were redrawn until no phone number appeared for more than one property. Appendix Table A1 shows that sampled and non-sampled properties are well-balanced with respect to the few characteristics available for all properties: there are no differences in property value and only minor differences in terms of property locations. Finally, selected properties are slightly more likely to be used for both domestic, and commercial or institutional purposes.

[26] The Transform Freetown Third Year Report can be found here. Information about all service provided between 2019 and 2021 can be found on pages 84 to 105.

[27] I excluded services with no fixed location, such as "the set up of youth tricycle groups to collect garbage" from the geo-referencing.

Table 1.2: Public services provided by the FCC

Category	Examples	#
Access to water	Installation of boreholes, water pumps, or milla tanks	137
Disaster management and prevention	Clearing or construction of drainages, or reconstruction support after floods and fires	53
Environmental management	Tree planting, or the construction of pedestrian crossings and street steps	32
Sanitation	Construction of public toilets, or the cleaning of illegal dumpsites	23
Market infrastructure	Repairs or improvements of market infrastructure such as roofs and floors	14
Health care	Infrastructural improvements of peripheral health units (PHUs) and hospitals	7
Education	Provision of furniture to schools, or the construction of early learning centers	6
Road maintenance	Road repairs, or the pavement of roads	5

Notes: This table provides an overview of services with a fixed location delivered by the Freetown City Council between 2019 and 2021. Column (1) illustrates the different categories of services provided. Column (2) gives specific examples of what has been provided. Column (3) shows how many services have been provided in each category. The table excludes services with no fixed location and services for which existence could not be verified.

vention, (3) environmental management, (4) sanitation, (5) the improvement of market infrastructure, (6) health care, (7) education, and (8) road maintenance. Table 1.2 gives an overview of these categories as well as a few examples of the specific services provided by the FCC for each category. Column (3) specifies how many services were provided between 2019 and 2021 in each category.

Figure 1.2 illustrates the geographical distribution of services provided by the Freetown City Council. While this map does not reflect overall public service provision in Freetown, it can be seen as a proxy for the FCC's current policy priorities in the area of service provision. Approximately 50% of all recorded projects fall in the category of water provision. As shown by the red squares, water projects are widely scattered throughout Freetown. They include the provision of water tanks or street taps, as well as the building of water wells. With roughly 19%, projects related to disaster management and prevention rank second. These types of projects, shown in dark blue, are more densely provided in the central parts of the city, and less so in the Western and Eastern area. Roughly 12% of all projects are related to environmental management. These services, marked in turquoise, are predominantly located at the Eastern part of the city center and in the Eastern outskirts of the city and include, among others, the clearing and building of drainage systems. With slightly above 8% and around 5%, sanitation and market infrastructure projects rank forth and fifth in the agenda of FCC service provision.

Finally, health care (2.5%), education (2%) and road maintenance (2%) are least represented in the provision of services by the FCC. The latter is at least partly due to the fact that the construction and maintenance of roads officially falls under the mandate of the federal rather than the local government.[28] Whereas this is also true for the provision of access to water—especially in terms of connecting neighborhoods and individual houses to the piping system—the FCC has made water a priority in the absence of sufficient action by the national government.

Figure 1.2: Locations of services provided by the local government
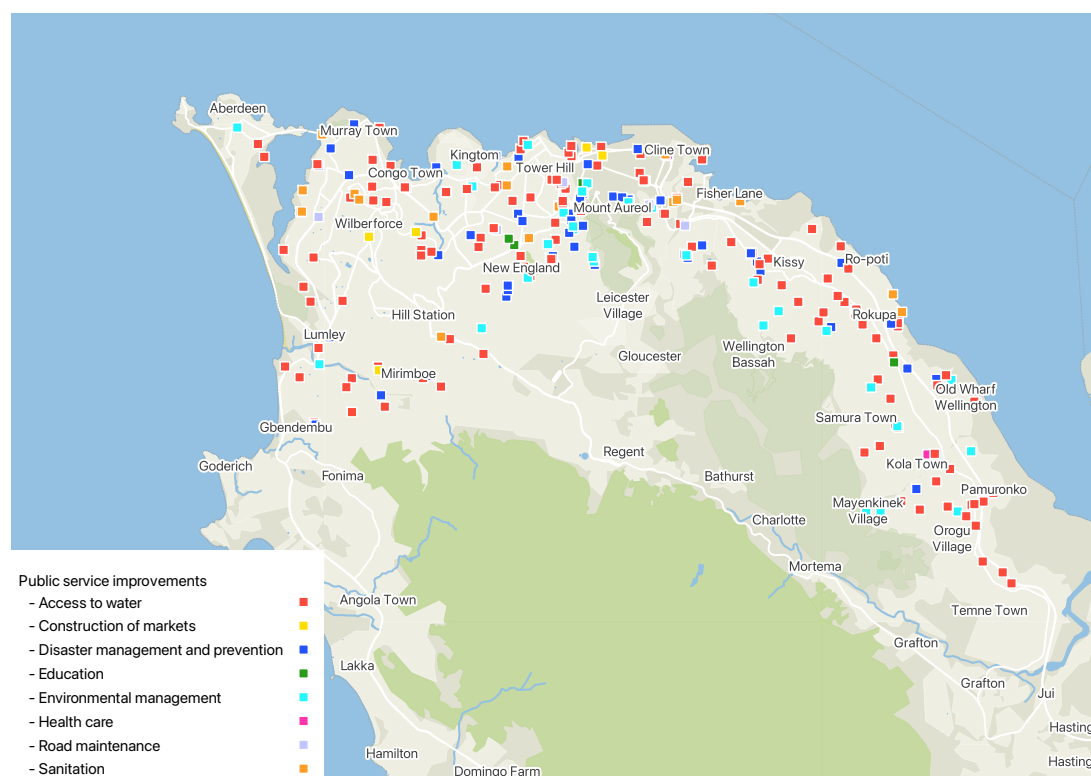
Notes: The figure illustrates the locations at which public services have been provided by the Freetown City Council between 2019 and 2021, color-coded by the type of service provided at each location.

Whereas many of the located services seemed to be well-functioning, others were rather short-lived. For instance, many of the drainages that had been cleared by the FCC at some point were again flooded with trash at the time of the geo-coding procedure. Overall, enumerators marked approximately 4% of all registered services as not currently functioning.[29]

---

[28] Note that the few road *construction* projects commissioned by the FCC were dropped from the list of services and only projects related to road *maintenance* were kept. This decision was taken to align the information provided to the treatment group (see Section 1.4.1) more carefully with the FCC's mandate.

[29] I did not exclude these services from the list of services I provided information about. However, only 2.5% of all treated individuals ended up receiving information about a service considered to be non-functional.

Survey data   During the baseline data collection, enumerators successfully conducted phone interviews with a total of 5,827 individuals.[30] Interviews were held with respondents who confirmed to live in or own the respective properties.[31] Given that effects on property tax payments were expected predominantly through an interaction with property owners, the latter were over-sampled.[32]

Participants were asked about their attitudes towards taxation and public service provision by the local government, and, importantly, about their preferences in terms of public service provision.[33] In particular, enumerators asked individuals to choose one out of eight areas of public service provision, which they would most want the FCC to provide services in and one area in which they would least want the FCC to provide services in.[34] The eight categories correspond to the previously identified types of services that had recently been provided by the FCC.

To decrease the likelihood of providing information about services that are irrelevant to the respondents due to their distance, individuals above the 97th percentile of distance to their most or their least preferred type of service were excluded from the study prior to randomization. Similarly, individuals with missing data on their most or least preferred public service were excluded. This leaves me with a final sample of 5,494 individuals. The sample is representative for the set of eligible properties in terms of property value, but contains a set of properties that is slightly more likely to be used for multiple purposes and different with respect to the geographical distribution of all eligible properties, suggesting selection into being surveyed based on observable characteristics. Balance results can be found in Appendix Table A2. Imbalances are accounted for in the analysis by controlling for both multiple use and the ward in which the property is located.

Approximately 8 to 12 days after the intervention, an endline survey including questions about attitudes related to taxation, public service provision and the local government more gen-

---

[30] Enumerators received daily call sheets with a random set of phone numbers to be called. In total, 11,985 phone numbers were attempted, with a response rate of 66%. Out of those individuals who were reached, 21% indicated that the neither lived in, nor owned the property they were called about, 5% did not consent to take part in an interview, and 73% consented. Due to uncertainties concerning the funding at this stage, it was not feasible to continue the data collection until the envisioned 7,000 individuals were reached.

[31] In total, 0.5% of respondents stated to own, but not to live in the respective property.

[32] This was done by splitting the data collection process into two batches and interviewing only owners in the second batch, while interviewing respondents without further selection in the first batch. The exact procedure is described in Appendix 1.A.

[33] A subset of individuals was also asked about their attitudes towards the local government more generally as well as their political participation. More information about the selection of these individuals can be found in Appendix 1.A.

[34] The exact framing of the survey questions was as follows: 1. In the following, I will read a list of areas of public service provision to you. Imagine the FCC was to provide services in only one of these areas. Which one would you personally prefer? 2. Now think about the same list of public service areas and imagine again that the FCC was to provide services in only one of these areas. Which one would you personally be least interested in?

erally was conducted. I use this data to get a better understanding of the potential mechanisms driving the effects on tax compliance. This survey was administered to a selected subset of the individuals interviewed at baseline among which property owners were not oversampled. As such, the endline sample is not fully representative of the overall study sample, in particular in terms of the share of owners.[35] To account for this, I conduct several adjustments which I explain in detail in Section 1.8.

Public service preferences   Figure 1.3 illustrates the public service preferences reported by the study participants at baseline. Shown in dark red is the percent of individuals who are most interested in the provision of a particular type of public service. The light red bars illustrate the types of services individuals are least interested in seeing provided. More than 50% of individuals declare water to be their top priority. This overlaps almost perfectly with the municipality's priority for providing water as presented in Table 1.2. Water is followed by a priority for roads and health care with 15.1% and 14.7%. Only 1.0% of individuals have a strong preference for environmental or disaster management, the FCC's second most important categories in terms of services provided. In fact, these two types of services rank first when it comes to least preferred services, followed by roads with 13.0%.

Elicited preferences can capture several layers of preference formation. First, individuals may select a specific type of service as their most preferred type because, all else equal, they value this type of service more than any other type of service. At the same time, when asked for the types of services *to be provided*, individuals' responses are likely informed not only by general preferences, but also by the baseline level of public service provision, or the perception thereof. If individuals care a lot about environmental services, but at the same time don't have access to water, they are likely to prioritize the provision of water over the provision of environmental services out of necessity. Thus, the measures of preferences used in this paper likely captures a mixture of general preferences for and perceived need of a particular service.

As I do not directly observe the overall level of services provided in the city of Freetown, I look at several alternative variables to get a better understanding of how stated preferences are formed. I use these variables to predict the likelihood of choosing a specific type of service as one's most or least preferred service in a seemingly unrelated regression framework. Preferences for a specific type of service are uncorrelated with distance to the closest service of that type, as well as with the amount of services recently provided within a 3 kilometer radius (see

---

[35] The share of owners in the endline sample is 64.4% as compared to 91.3% among those who did not participate in the endline survey. In addition, endline participants are on average somewhat younger, live further away from the closest location at which a service of their most preferred type was provided by the FCC, and are less likely to have a preference for the provision of water. Finally, they live in properties with somewhat lower value. See Appendix Table A3 for more details. The sampling procedure is described in more detail in Appendix 1.A.

Figure 1.3: Public service preferences at baseline
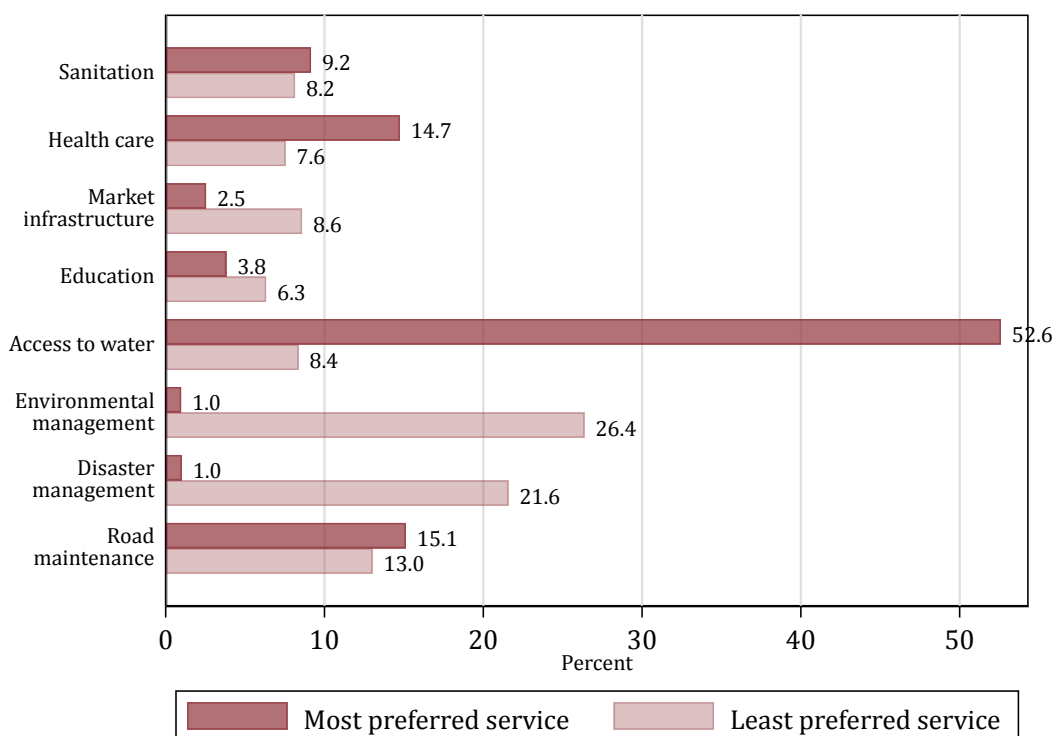
Notes: The figure illustrates the percentage of individuals indicating a specific type of service as their most (dark red) or least (light red) preferred service, respectively. The data used stems from the baseline survey and includes responses from all individuals included in the final sample.

Appendix Figure A4).[36] However, the lower perceived availability of or satisfaction with the provision of a specific type of service, the more likely individuals are to indicate this service as their most preferred one (see Appendix Figure A5). This suggests that preferences are indeed formed to a certain extent on the basis of perceived needs. Finally, preferences also vary by property value and property ownership.[37] To account for the fact that preferences correlate with specific characteristics of properties and their residents, I control for the above-mentioned variables in all regressions aiming at disentangling the treatment mechanisms.

---

[36] Results are similar when restricting to a 1 kilometer radius or extending to a 5km radius.

[37] In particular, higher property value is associated with a higher likelihood of having a preference for the provision of water. Property owners are significantly more likely to have a preference for roads than individuals who do not own the property they live in.

1.4 Experimental Design and Randomization

1.4.1 Experimental Design

I design and study two interventions disseminating information about public services provided by the Freetown City Council (FCC), a preference *match* and a preference *mismatch* intervention. To do so, I leverage the previously described geo-location of public services as well as the collected information about individual preferences. The information interventions were carried out as a combination of a call and a follow up short message (SMS) sent on the respondent's phone immediately after the call, and aimed at increasing awareness about municipality level public service provision. During the phone call, individuals were informed about one of the eight types of services identified above, including an example location at which this type of service had been provided, as well as the approximate commuting time from the respondent's home.[38]

The type of service mentioned during the call varied by individual depending on their preferences for public service provision elicited at baseline. In the *match* intervention (T_M), individuals received information about their most preferred service, whereas individuals randomized to receive the *mismatch* intervention (T_L) received information about their least preferred service. To target the information treatment geographically, individuals were informed about the closest location at which a specific type of service had recently been provided.

Individuals in the control group (Control or C) did not receive any information call or SMS. This pure control group does not allow me to directly assess whether the effects of my treatment are actually driven by the fiscal exchange content of the message, or whether individuals react to the treatment based on an alternative channel.[39] Thus, to shed light on the role of potential alternative mechanisms, I present suggestive evidence from heterogeneity analyses as well as analyses of the endline data in Section 1.9.

Appendix Figure A6 depicts the script used for both the *match* and the *mismatch* intervention. The two scripts differ in only one sentence, as can be seen when comparing Figure A6a against Figure A6b. Whereas the *match* group was reminded of their baseline preference in order to make the match more salient, the *mismatch* group was simply informed that some of the FCC's budget went into providing a certain type of service.[40] Immediately after the phone call,

---

[38] To provide information on the commuting time, I made use of two datasets. First, I used the geo-locations of properties available in the tax registry. Then, I used the geo-mapped locations at which public services had been provided by the FCC and generated approximate travel times based on the distance between the closest service of each type and the respondent's home.

[39] Note that due to budgetary constraints, it was unfortunately not possible to include an additional control group receiving a non-fiscal exchange message in the study.

[40] To avoid negative associations with the phone call, individuals in the mismatch group were not specifically told that they received information about what they previously mentioned to be their least preferred good.

enumerators sent an automated SMS to the respondents. This SMS summarized the most important information given during the phone call, i.e., the type and location of a service provided by the FCC, as well as the individual's approximate commuting time.[41]

While the phone call script differs only insubstantially across treatment arms, the content of the treatment message varies depending on individual preferences and the location of the property in which individuals live. Given that the distribution of most and least preferred services differs (see Figure 1.3), so does the distribution of the types of services individuals are informed about in the match versus mismatch treatment (see Appendix Figure A8).[42] Overall, roughly 30% of individuals were informed about the provision of access to water, the vast majority of which are individuals in the match group. Information about water makes up more than half (53.9%) of the match messages. Road maintenance and health care rank second and third for the match group. Whereas environmental management and disaster management rank last with 0.8% each in the match group, these two categories rank first and second in the mismatch group.

These differences illustrate well that there is non-random variation in the types of services individuals are informed about across treatment groups. Similarly, conditional on being informed about the same type of service, distance to services differs across treatment arms for some services.[43] The implications of this variation for the interpretation of treatment effects are discussed in detail in Section 1.6.

1.4.2 Randomization

The treatment was randomly assigned on the individual level. Among all 5,494 study participants, one third (1,834) of individuals was randomized into the *control group* (C), one third (1,829) into the *match treatment* (T_M), and one third (1,831) into the *mismatch treatment* (T_L), respectively. The randomization was done using stratification by gender, education, ownership status, and geography, as well as the survey batch and a dummy variable indicating a baseline preference for road maintenance.[44] The latter were included to guarantee balance across treatment groups with respect to service preferences that do not reflect the local government's

---

[41] An example of such message is shown in Appendix Figure A7.

[42] Figure A9 in the Appendix shows the share of individuals receiving information about a certain type of service independent of the treatment group assignment.

[43] In particular, individuals in the match group who received information about sanitation are significantly closer to the nearest location at which sanitation services have been provided than individuals in the mismatch group who received information about sanitation. Similarly, individuals in the mismatch group informed about services related to market infrastructure and road maintenance are significantly closer to these services than individuals in the match group informed about the same types of services. There are no significant differences in distance to the respective service between the match and mismatch group for the remaining types of services.

[44] Survey batches refer to whether individuals where in the random respondent or owner batch. Whereas the former were to be interviewed again at endline to measure self-reported outcomes, the latter were included to enlarge the sample used to measure actual tax compliance. More details on the two batches can be found in Appendix 1.A.

mandate or priorities. The randomization was successful in generating a well-balanced sample with respect to the main baseline characteristics, with minor imbalances in terms of respondent age (see Appendix Table A4).[45] Importantly, individuals are also well-balanced with respect to baseline preferences and perfectly balanced in terms of distance to the closest location at which services of each type were provided by the FCC (see Appendix Table A6.)

1.5   Fiscal Exchange: Theory and Hypotheses

Fiscal exchange refers to the concept of a reciprocal relationship between the government and the taxpayer, where individuals are willing "to pay taxes in exchange for benefits that the state provides to them or to others even though their pecuniary payoff would be higher if they didn't pay taxes" (Luttmer and Singhal, 2014, p. 150). The stronger this relationship, the more likely they are thought to pay taxes. But what defines the (perceived) strength of the fiscal social contract and how can public service messages affect it?

A crucial dimension of fiscal exchange is the individual's own perceived benefit of public service provision: specific services can be relevant for some individuals or households, but not for others. Several papers have pointed to the importance of this channel. Evidence from the lab suggests that individuals are less likely to evade taxes when tax revenue is allocated to programs they support (Alm et al., 1993). Similarly, in a field experiment, Giaccobasso et al. (2023) shows that information about the share of property tax revenue used to finance public schools affects the probability of filing a tax appeal differentially for households with children, who directly benefit from schooling expenditures, and households without children, who do not benefit directly.

These examples illustrate that the extent to which individuals directly (including their household members) benefit from public service provision can affect tax-related outcomes. It also crucially demonstrates that personal benefit can have various dimensions. Individual level preferences as well as the accessibility of services determine how beneficial public service provision is perceived to be. Furthermore, depending on the specific context, certain types of services may be perceived to be more essential than others and can thus play a role independent of stated individual-level preferences for public service provision.

The design of my experiment allows me to look at the role of individual preferences, accessibility, and contextual importance of a specific service separately by investigating how individuals react to the treatment depending on the exact content of the message. Through the match and mismatch treatment, information about public service provision is either aligned or

---

[45] Table A5 in the Appendix further shows that balance is also given for all characteristics but age and water being one's least preferred service for the subsample of individuals who received an endline survey, which is used to look at self-reported outcomes.

misaligned with personal preferences. Information on accessibility with respect to the location of the service is given through the approximate commuting time between the location of the service and the individual's house. Finally, the type of service individuals are informed about can be considered independently of individual-level preferences as reflective of the more general needs of the population.

In any setting, and in particular in the context of Freetown, access to water can be considered a basic need without which many other public services have less value. For instance, without access to water, health facilities cannot fully operate. In Freetown, water access is still far from universal. Evidence from a pilot study in a subset of wards suggests that around 39% of registered properties have no access to water at all, and only around 45% are connected to the piped water system (Grieco et al., 2019). While this data is not representative for the entire city, it is important to mention that these numbers do not take informal settlements into consideration, where the situation is even more dire.

In response to the under-provision of access to water through the federal government—which is officially responsible for the provision of water, and in particular the piped water system—the Freetown City Council has increased its efforts to provide water to the citizens of Freetown. This is clearly visible when considering the share of services provided with respect to access to water: between 2019 and 2021, 49.5% of all services provided were related to water (see Table 1.2). At the same time, while many individuals state other preferences, the majority of individuals interviewed for this project considers increased access to water to be the top priority for municipality level public service provision (see Figure 1.3). Taken together, these facts underline the crucial importance of increasing access to water in Freetown.

In line with this, the main part of this paper focuses on disentangling the role of three potential indicators of perceived personal benefit. To do so, I look at variation in the information provided to individuals during the treatment call along the following dimensions: (1) the alignment with personal preferences (*match* vs. *mismatch* group), (2) distance of the service, and (3) whether individuals were informed about services providing access to water or any other type of service.

Given that more personal benefit is considered to positively affect individuals' tax morale, I expect individuals to increase tax compliance if they receive information about (a) their most preferred service, (b) a service considered to be accessible with respect to its location (i.e., a service considered to be close), or (c) a service that is unambiguously considered to be important independent of own stated preferences (i.e., water). Tax payments are expected to increase in the share of information aligned with personal benefit. At the same time, it is ex ante unclear whether information that does not suggest personal benefit has no or even negative effects on compliance. This question will be answered empirically in the Section 1.7.

1.6   Measurement and Estimation

This article examines the effects of information about public service provision on tax compliance using administrative data. A number of studies look at declarations of income or revenue, which determine the tax base, but not necessarily true compliance (e.g., Bott et al., 2020; Mascagni and Nell, 2022). By contrast, administrative data allows me to capture actual tax payments made. I predominantly look at tax payments made for the 2022 tax cycle between April and September, 2022, i.e., after the intervention and up until the official tax payment deadline for the 2022 tax cycle. To further examine the longevity of the treatment effects, I additionally consider payments made within the entire 2022 tax cycle, i.e., until December 2022. I define three main outcomes: (1) a variable measuring the sum of all payments made by a certain point in time, (2) a dummy variable indicating whether any payment was made, and (3) a variable measuring the sum of all payments made conditional on any payment.

I use OLS to estimate the following equation for the average treatment effect of receiving information about public services provided by the FCC :

$$y_{ik} = \beta_1 T_{ik} + \alpha_k + \mathbf{X}_{ik}\mathbf{\Gamma} + \epsilon_{ik}, \tag{1.1}$$

where i indexes individuals and k the randomization strata. T_{ik} is an indicator for individuals that received any information treatment. β_1 estimates the average causal effect of assignment to any treatment on the respective outcome of interest, y_{ik}. α_k are strata fixed effects and $\mathbf{X}_{ik}$ are individual-level covariates. In the specifications presented in the main part of this paper, these include the baseline value of the outcome, y_{ik0}, as well as baseline enumerator fixed effects. In addition, I control for all variables which are either unbalanced across treatment groups or unbalanced with respect to the set of properties the sample is expected to represent. I therefore include respondent age, ward fixed effects, and a variable indicating whether a property is used for commercial or institutional purposes on top of domestic ones. When looking at self-reported endline outcomes, I further include enumerator fixed effects for the endline interview and a dummy variable indicating the timing of the endline data collection (first vs second week). All regressions are estimated using robust standard errors. Finally, I use inverse probability weights to adjust for the fact that the probability of assignment to a treatment group or the control group varies slightly across strata.

To look at the effects of the two treatment arms separately, I estimate

$$y_{ik} = \beta_2 T_{Mik} + \beta_3 T_{Lik} + \alpha_k + \mathbf{X}_{ik}\mathbf{\Gamma} + \epsilon_{ik}, \tag{1.2}$$

where T_{Mik} and T_{Lik} refer to the *match* and the *mismatch* treatment, respectively, and β_2 and β_3 estimate the effects of being assigned to either of the two treatment groups on the respective

outcome of interest, y_{ik}.

Whereas the treatment arms were designed to causally investigate the role of public service preferences, non-random variation in the type of service that individuals were informed about, as well as the distance to that service complicates the interpretation of the results. Thanks to the randomized assignment, individuals are balanced in terms of preferences across treatment groups. At the same time, the share of individuals who mention a specific type of service as their most preferred service differs substantially from the share of individuals mentioning the same service as their least preferred service. Thus, the distribution of services individuals were informed about in the two treatment groups differs, and with it distance. Potential differences in effect sizes across the treatment arms could therefore be driven by differences in the reaction to information on specific types of services or services within a certain range from the individual's property.

To account for this, I do two things. First, I introduce additional control variables which may affect the formation of preferences (see Section 1.3). Second, I exploit the variation in the treatment message to look at heterogeneity depending on the type of service and distance to the service individuals are informed about. This helps me to separate the role of service preferences from that of distance and service type, and thus allows me to disentangle several potential drivers of perceived personal benefit.

1.7 Main Results

1.7.1 Average Treatment Effects on Tax Compliance

This section presents the effects of providing specific, targeted information about local level public service provision on tax compliance. I begin by illustrating the effect of receiving any type of information about local level public service provision on three main tax compliance outcomes: (1) the sum of all payments made, (2) a dummy indicating whether any payment was made, and (3) the sum of all payments made conditional on payment. All three outcomes are measured at the official tax payment deadline end of September 2022, approximately six months after the intervention.

Column (1) of Table 1.3 shows that the treatment significantly increases timely tax payments by SLE 23.3 (USD 1.5), an increase of 20% as compared to the control group mean.[46] This suggests that targeted information about public services provided by the institution responsible for collecting taxes can indeed have positive effects on tax compliance in a time frame crucial for the tax authorities. With 3,660 treated individuals, the treatment increased tax payments before

---

[46] The effect is robust to winsorizing payments at the 95th percentile of the distribution of the share of tax paid, though somewhat smaller in size (see Appendix Table B7).

Table 1.3: Treatment effects on tax compliance by the payment deadline (September 2022)

	(1) Amount paid (SLE)	(2) Paid any tax (d)	(3) Amount paid (SLE, cond.)
Treated	23.784**	0.012	74.983
	(11.322)	(0.008)	(55.252)
Obs.	5,494	5,494	1,159
Control mean	118.829	0.201	590.605
SD	418.244	0.401	769.365

Notes: Robust standard errors in parentheses. Significance levels are indicated by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Regressions include strata, enumerator, and ward fixed effects and control for respondent age, multiple property use, and the respective outcome variable at baseline.

the deadline by more than SLE 87,000, or approximately USD 5,600.[47] Extending the intervention to all 46,844 eligible properties could increase tax revenue by more than USD 71,700. With a conservative per-individual cost of SLE 14.2 (including the cost of mapping all services as well as enumerator and phone costs for implementing the intervention), the intervention raises 1.7x its cost. Given that almost 50% of the cost goes back to phone costs and the government often reaches more advantageous agreements with network providers, this figure can be interpreted as a lower bound for the cost-effectiveness of providing public service information at scale.[48]

The effect on overall revenue could be driven by several reactions to the treatment: an increase in the likelihood of any tax being paid (extensive margin), an increase in the amount paid by those who do pay (intensive margin), or a combination of the two. While neither the extensive (Column 2) nor the intensive margin effects (Column 3) are significantly different from zero on their own, both coefficients are positive and sizable. This suggests that the treatment may incentivize some individuals to switch from not paying to paying taxes, while it encourages others to pay more than they otherwise would have paid.[49]

To get a more nuanced understanding of how individuals react to the treatment, I look at two different types of individuals: poor individuals, who reside in low-value properties, and rich individuals, who reside in high-value properties. These individuals vary along several dimensions. By definition, individuals residing in high-value properties face higher tax liabilities in absolute terms. They are also more likely to pay taxes in the absence of treatment.[50] In ad-

---

[47] Nominal exchange rates for September 30, 2022 taken from https://www.exchangerates.org.uk/USD-SLL-spot-exchange-rates-history-2022.html.

[48] In previous interventions conducted by the government itself, network providers even allowed free usage of their network.

[49] Note that individuals in the control group who did pay taxes by the end of September paid on average 90.5% of their outstanding tax liability.

[50] This could be driven by the positive relationship between property value and the probability of experiencing en-

dition, poor and rich individuals vary in their access to and reliance on public services. As the assessed value of a property depends, among other things, on access to public services, residents of low-value properties have on average lower access to public services. At the same time, they are on average poorer, and as such more budget-constraint. Thus, they depend quite heavily on the provision of public services. Residents of high-value properties, on the other hand, often provide services privately and as such depend less on the government's fulfillment of the social contract.

I look at the variation in treatment effects empirically by splitting my sample into properties with below (low-value) and properties with above (high-value) median value. I present the effects on tax compliance separately for both subsamples. Columns (1) and (2) of Table 1.4 show that tax payments increase significantly for both below and above median value properties, the effect being significantly larger among the latter. This is not surprising given that these properties have on average tax liabilities that are more than five times as high as those of below median value properties.[51] At the same time, it is not obvious that individuals living in high-value properties, who have on average better access to public services and are more likely to provide services privately, would react to a fiscal exchange message at all. Understanding the dynamics behind these effects thus merits further investigation.

Table 1.4: Treatment effects on tax compliance by property value (September 2022)

	Amount paid (SLE)		Paid any tax (d)		Amount paid (SLE, cond.)	
	(1)	(2)	(3)	(4)	(5)	(6)
	< median	> median	< median	> median	< median	> median
Treated	3.861**	47.718*	0.012	0.009	8.554	51.419
	(1.785)	(24.763)	(0.010)	(0.013)	(7.904)	(96.423)
p-val T S1 = S2	0.054		0.486		0.610	
Obs.	2,747	2,747	2,747	2,747	446	713
Control mean	23.681	209.321	0.153	0.247	154.533	848.112
SD	60.846	566.688	0.360	0.431	62.768	872.514

Notes: Robust standard errors in parentheses. Significance levels are indicated by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Regressions include strata, enumerator, and ward fixed effects and control for respondent age, multiple property use, and the respective outcome variable at baseline. p-values below the treatment effect show whether there are significant differences between the treatment effects in sample 1 (below median property value) and sample 2 (above median property value).

Columns (3) – (6) illustrate the extensive and intensive margin effects separately. Again, while not significant on their own, both effects seem to play a role. The point estimates suggest that the treatment mainly incentivizes residents of lower-value properties to pay taxes. As com-

---

forcement. It could also be explained by individuals residing in low-value properties being more cash-constraint than those living in high-value properties.

[51] This is true both for annual and outstanding (incl. arrears and penalties) tax liabilities.

pared to the control group mean of 15.3%, they are up to 8% more likely to pay taxes. Among above median value properties, who are around 61% more likely to pay taxes in the absence of any treatment, the effect would be less than half that size. As compared to this, the coefficients on the amount paid conditional on any tax paid are similar in relative size, but much larger for high-value properties in absolute terms.

Based on these results, I want to gain a better understanding of the potential mechanisms driving different reactions to the treatment. I start by looking at the role of service preferences measured through the two different treatment arms. To compare the effects of the two treatment arms for individuals with identical preferences, I add service type fixed effects indicating individuals' most and least preferred service. To further account for the fact that preferences are formed endogenously, I control for several variables that are likely to affect the formation of preferences. In particular, I include distance to the closest location at which the most and least preferred services are provided, perceived availability of the most and least preferred service at baseline, property value, and property ownership.

Table 1.5: Treatment effects on tax compliance by treatment arm (September 2022)

	(1) Amount paid (SLE)	(2) Paid any tax (d)	(3) Amount paid (SLE, cond.)
Match	7.894	0.012	13.145
	(9.857)	(0.009)	(43.045)
Mismatch	35.413**	0.015	143.640**
	(14.277)	(0.009)	(56.784)
p-val. Match = Mismatch	0.057	0.778	0.011
Obs.	5,469	5,469	1,151
Control mean	118.829	0.201	590.605
SD	418.244	0.401	769.365

Notes: Robust standard errors in parentheses. Significance levels are indicated by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Regressions include strata, enumerator, and ward fixed effects and control for respondent age, multiple property use, and the respective outcome variable at baseline. In addition, regressions control for public service preferences (most, least), linear distance to closest services (most, least), perceived availability of services at baseline (most, least), property value, and property ownership.

Table 1.5 presents the effects on tax compliance by treatment arm. Differently than expected, the average treatment effect seems to be driven by individuals who receive information about their least preferred service.[52] The effects on the amount paid, both unconditional and

---

[52] Appendix Tables B1 to B3 show how the effects change with the gradual inclusion of these control variables. The effects are similar across specifications, with the effects of the mismatch treatment becoming significantly different from those of the match treatment only once controlling for type of service, distance, and perceived availability at baseline. The results are also robust to replacing linear distance to services controls with indicators for type-specific above median distance, distance quartiles, or linear distance to the closest service of each type (see Appendix Tables B4 to B6).

conditional on payment, are significant and differ from the effects of the match treatment. This has two important implications. First, in this setting, there is no backlash effect of receiving potentially negative information about the provision of a service previously indicated to be one's least preferred service. Second, there has to be another dimension explaining why, on average, individuals react to the mismatch, but not the match treatment.

To get a more differentiated view, I once more look at the effects for below and above median value properties separately. Table 1.6 offers a new perspective, showing that the large and significantly different effects of the mismatch treatment are driven by individuals residing in properties with above median value. Residents of below median value properties, on the other hand, seem in fact to react (more) to the match treatment. These results indicate that the anticipated fiscal exchange channel may indeed be more relevant for individuals who are more dependent on the provision of services by the government. At the same time, the coefficient on the mismatch treatment is still positive and insignificantly different from the coefficient on the match treatment, suggesting that the value of targeting information to individual preferences may be lower than anticipated based on previous research (Alm et al., 1993; Lamberton et al., 2018; Giaccobasso et al., 2023). This finding is likely related to context-specific factors such as the overall level of public service provision and the related absolute valuation of services provided, independent of their ranking. Importantly, in low-capacity settings, where public service provision is generally insufficient, specific information about any type of service, independent of preferences, may be considered valuable and incentivize a positive reaction towards the state.

These results improve our understanding of the role of different types of information taking both preferences and the allocation of public services within a setting as given. However, individuals receive—by definition of the preference matching—information about different types of services which are located differently close to their homes. These factors can be indicative of how much individuals benefit from a specific service independently, and they are likely to interact with the effect of matching preferences. To investigate whether that is indeed the case, the following section looks at the role of distance and type of service, and their interaction with public service preferences, in more detail.

1.7.2 The Components of the Treatment Message

To disentangle the different dimensions along which perceived personal benefit of public service provision can affect tax compliance, I look at three components of the treatment message: alignment with individual preferences, distance, and type of service. As before, I look at the effects of the *match* and the *mismatch* treatment to investigate the role of service preferences. To examine the role of information about the provision of access to water, I introduce two dummy variables indicating whether water is an individual's most or least preferred service,

Table 1.6: Treatment effects on tax compliance by property value and treatment arm (September 2022)

	Amount paid (SLE)		Paid any tax (d)		Amount paid (SLE, cond.)	
	(1) < median	(2) > median	(3) < median	(4) > median	(5) < median	(6) > median
Match	5.156**	10.153	0.018	0.004	12.238	-51.891
	(2.263)	(20.812)	(0.012)	(0.015)	(9.271)	(85.179)
Mismatch	3.379	70.854**	0.012	0.013	0.625	164.088*
	(2.139)	(29.996)	(0.012)	(0.015)	(11.049)	(89.828)
p-val. Match = Mismatch	0.465	0.038	0.661	0.578	0.334	0.015
p-val. Match S1 = S2		0.290		0.666		0.808
p-val. Mismatch S1 = S2		0.046		0.456		0.299
Obs.	2,736	2,733	2,736	2,733	444	707
Control mean	23.681	209.321	0.153	0.247	154.533	848.112
SD	60.846	566.688	0.360	0.431	62.768	872.514

Notes: Robust standard errors in parentheses. Significance levels are indicated by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Regressions include strata, enumerator, and ward fixed effects and control for respondent age, multiple property use, and the respective outcome variable at baseline. In addition, regressions control for public service preferences (most, least), linear distance to closest services (most, least), perceived availability of services at baseline (most, least), property value, and property ownership.

respectively.

Finally, to assess the role of differences in access to services, I introduce an indicator for distance to services. By virtue of the intervention's design, treated individuals received information about services with varying distance to their homes. Whereas at the minimum, individuals live less than 10 meters away from the treatment service, at the maximum this distance reaches almost 9 kilometers.[53] This suggests that, despite the intention to provide geographically relevant information to everyone, accessibility of these services differs largely across individuals—potentially affecting the extent to which information about service investments positively affects tax compliance.

To account for this, I define binary variables for distance to the most and least preferred type of service using below and above median distance to the respective service. Distance to services not only depends on the property location and the local government's prioritization for public service investments, but also inherently on the type of service. To take this into consideration, I define median distance using service-specific distances rather than the overall distribution of distance to the treatment service independent of its type.[54]

To look at the role of all three dimensions of perceived personal benefit, I interact the match treatment with the indicator for above median distance to the most preferred service and the indicator for water being the most preferred service. Similarly, I interact the mismatch treatment with the indicator for above median distance to the least preferred service and the indicator for water being the least preferred service. In addition, as in the previous specifications, I include service preference fixed effects, linear distance to the most and least preferred service, perceived service availability at baseline, property value, and ownership status as controls.

Figure 1.4 illustrates the interacted effects on whether any tax was paid by end of September for below and above median value properties graphically. The markers show treatment effects with 90% confidence intervals for each treatment arm (match vs. mismatch), differentiating by service accessibility ($<$ vs. $>$ median distance) and type of service (water vs. other type). While there are no significant differences in effect sizes depending on the content of the message for above median value properties, an interesting picture emerges for below median value properties. In line with the hypotheses formulated in Section 1.5, the point estimates are larger when

---

[53] Figure B1 in the Appendix shows the distribution of distance to the treatment service. By definition, the graph is based on data from individuals in the treatment group only, since individuals in the control group did not receive any information about service provision.

[54] Specifically, I look at the sample distribution of distance to the closest location at which a specific type of service is provided. I then define the median distance for each type of service. This allows me to categorize the minimum distance between each property and each type of service into below or above the median. Figure B2 in the Appendix illustrates the distribution of distance for each type of service. Depending on the type of service an individual received information about, I replace the indicator variable for below or above median distance to the treatment service with the property-specific indicator for that type of service. I then run regressions excluding either treated individuals with below or above median distance to the treatment service.

Figure 1.4: Heterogeneity of the treatment effect on whether any tax was paid depending on the content of the treatment message (September 2022)
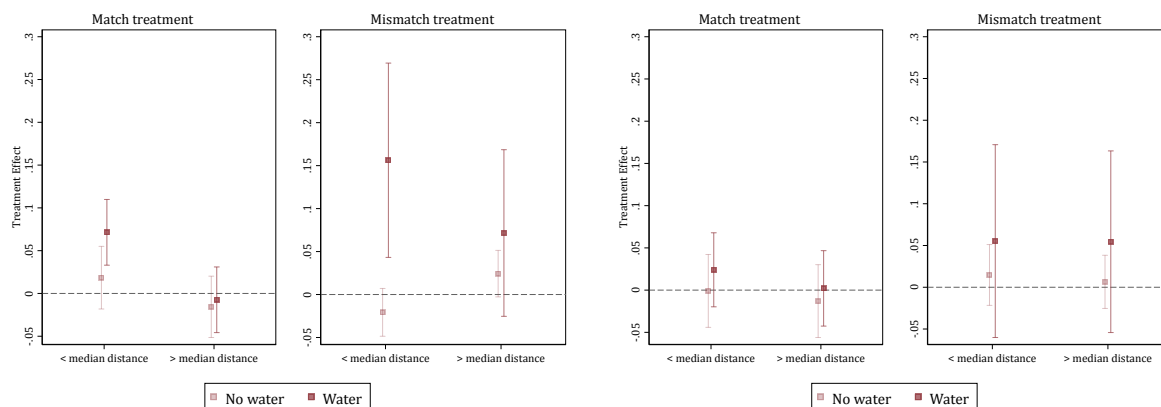
(a) Below median property value   (b) Above median property value

Notes: The figure illustrates the effect of being treated depending on whether an individual received the match or the mismatch treatment, whether thy received information about water or another type of service, and whether the respective service has below or above median distance. Figures 1.4a and 1.4b show effects separately for low- and high-value properties. Coefficients are shown with 90% confidence intervals. All regressions include strata, enumerator, and ward fixed effects and control for respondent age, multiple property use, and the respective outcome variable at baseline. In addition, regressions control for public service preferences (most, least), linear distance to closest services (most, least), perceived availability of services at baseline (most, least), property value, and property ownership. Appendix Figure B4 illustrates the effects for the entire sample.

individuals are informed about water as compared to any other type of service. This is true both for the match and the mismatch treatment. The point estimates are also larger in three out of four cases when information about a service at below median distance is provided.[55]

Overall, the likelihood of paying taxes among lower-value properties is affected significantly only when the treatment message informs about water at below median distance. When such information is provided to individuals with a preference for the provision of access to water, the likelihood of paying taxes increases by 7.1 percentage points or 62.1% as compared to individuals with a preference for water and below median distance to water in the control group, of which 11.5% paid taxes by the deadline. For individuals who receive the information about water at below median distance, but consider access to water their lowest priority, the effect amounts to 15.6 percentage points or 134.4% as compared to their control group counterpart, which has a compliance rate of 11.6%.[56] Both effects are significantly larger than the respective effects of the match or mismatch treatment when information about another type of service at

---

[55] Appendix Figure B4 shows that these patterns do not extend to the intensive margin, i.e., to effects on the amount of tax paid conditional on payment.

[56] The effects on individuals informed about water at below median distance, which seem to be driven by below median value properties, are visible, although somewhat less pronounced, when looking at the entire sample as well (see Appendix Figure B3).

below or above median distance is provided. For the match treatment, the effect is also significantly larger when information about water at below as compared to above median distance is provided.

While the effect appears to be larger for individuals in the mismatch group, the difference in coefficients between the match and mismatch group is not statistically significant. The lack of precision in the estimated effect of information about water misaligned with preferences is likely related to the small number of individuals who state water as their lowest preference, have below median distance to water, and are part of the mismatch treatment arm. Among below median value properties, only 34 individuals in the mismatch group receive information about water at below median distance, whereas this number amounts to 226 individuals in the match group.

A potential explanation for differences in the size of treatment effects across treatment arms could be that individuals are differentially informed about the existence of these services prior to the intervention. Differences in the level of prior information may create differences in the perceived value of the provided information, with less informed individuals perceiving the quantity of information received to be higher. Combined with the overall value of information about water with below median distance, this could explain why individuals in the mismatch group show a potentially larger reaction to the treatment than individuals in the match group. Comparing treated individuals who report water to be their least preferred service to those who indicate water to be their most preferred service indeed indicates that the former are somewhat less likely to have known about the service before the treatment call. However, these differences are—likely due to the comparably small number of observations—not significant.

These findings show that specific information about public services provided by the local government has the potential to substantially increase tax compliance. Crucially, the results underline the importance of distance to the locations at which services are provided, i.e., the importance of service accessibility (on a geographical level). They also show that information on certain types of services is valued more than others, even if these services do not match stated individual preferences. These results underline the importance of considering perceived personal benefit of public service provision along multiple dimensions.

The extensive margin effects associated with high personal benefit are driven by below median value properties. Residents of these properties have on average lower access to public services and significantly lower monthly income. This suggests that individuals living in lower-value properties rely on average more on services provided by the local government. It is therefore to be expected that these individuals are particularly sensitive to information on services they personally benefit from. As opposed to this, individuals living in higher value properties do not respond differentially to the specific content of the treatment message, neither on the extensive nor the intensive margin. For these properties, the observed effects seem to be driven

by a more general fiscal exchange channel, with specific information about any public service independent of personal benefit signaling the government's capacity and willingness to fulfill the social contract. I investigate this channel in the following subsection.

Robustness of Main Results   The results for the average effect of being treated, as well as the message-specific effects on whether any tax payment was made are robust to a variety of alternative specifications, including specifications controlling for respondent baseline characteristics, specifications controlling for potential correlates of individual level preferences, and specifications without weight-adjustments for the probability of being treated, as well as specifications excluding control variables (see Appendix Tables B7 and B8). Importantly, the results are also robust to winsorizing the two variables indicating the sum of conditional and unconditional payments made at the 95th percentile of the distribution of the share of outstanding liability paid. Finally, the results also hold true when adjusting for the oversampling of property owners.[57] This finding is of particular importance for a potential scale up, as it proxies the effect sizes to be expected when information is given to whomever is reached over the phone, independent of property ownership.

1.8   Perceived Government Capacity and Service Provision

Perceived fiscal exchange may not only be determined by the extent to which individuals personally benefit from public service provision. Receiving information about the government's efforts to increase service delivery may also improve individuals' attitudes towards the government—including perceptions of its capacity to provide public services and fairness of tax rates—and through this the willingness to pay taxes (e.g., Hofmann et al., 2008; Slemrod, 2007). The treatment may also induce individuals to update their beliefs about the provision of public services more generally, increasing the perceived availability of or satisfaction with public service provision. In line with this hypothesis, previous studies have found positive correlations between satisfaction with public services and willingness to pay taxes or the belief in the citizens' obligation to pay taxes (e.g., Van den Boogaard et al., 2020; OECD, 2019; Bodea and LeBas, 2016; Hanousek and Palda, 2004). I investigate these mechanisms using self-reported data on perceptions related to service delivery and local government performance more generally collected during baseline and endline interviews. By design, I can link all self-reported information to administrative tax outcomes and thus examine directly how these dimensions interact with the treatment effect and thus tax compliance.

---

[57] In particular, this specification uses probability weights adjusting the weight of each property owner in the overall sample to reflect the share of owners achieved in the first part of the baseline survey when owners and tenants were interviewed randomly depending on who picked up the phone.

Variables of Interest   To investigate the treatment's effect on attitudes towards the government, I examine citizens' beliefs about the performance of the local government, as well as its responsiveness. I follow Weigel (2020) in the definition of the first outcome. *Performance of local government* is an index reflecting citizens' trust in and approval of the local government more generally (Besley and Persson, 2009).[58] *Local government responsiveness* is a variable increasing in the perceived responsiveness of the local government to its citizens' needs.

Second, to better understand the role of perceived public service provision, I look at a variety of indices reflecting individuals' perceived availability of and satisfaction with public services provided by the local government:

  (i) *Perceived availability of public services*: an index combining questions on perceived availability of each of the eight types of services (sanitation, health care, the improvement of market infrastructure, education, access to water, environmental management, disaster management and prevention, and road construction and maintenance) on 3-point Likert scales. These questions refer to the amount of provided services independent of quality.

 (ii) *Perceived availability of most preferred service* and *Availability of least preferred service*: standardized indicators reflecting survey answers to the questions on perceived availability of the individual's most and least preferred service, respectively.

(iii) *Satisfaction with public services*: an index combining questions on satisfaction with each of the eight types of services on 5-point Likert scales. These questions refer to the quality and adequacy of the amount of services provided.

(iv) *Satisfaction with most preferred service* and *Satisfaction with least preferred service*: standardized indicators reflecting survey answers to the questions on satisfaction with the individual's most and least preferred service, respectively.

In addition, I look at heterogeneous treatment effects by perceived availability of public services at baseline. Baseline perceptions of public service provision by the local government are very negative. 44.2% of individuals report that they think the local government provides no services at all in accordance with their most preferred type, and only 3.8% report that there are a lot of services aligned with their priority. These numbers are similar, although a bit more positive for the least preferred service, with 37.6% of individuals perceiving there to be no services and 10.8% perceiving there to be lots of services aligned with their least preferred

---

[58] I construct all indices by standardizing each survey question and taking the unweighted average of all index components. I use these indices as a means to reduce type I and II errors stemming from multiple hypothesis testing.

service type. To investigate how these attitudes affect the response to the treatment message, I interact a dummy variable indicating whether perceived availability of the most preferred service at baseline is low (no services) or high (a little or a lot of services) with the match treatment. Similarly, I interact the mismatch treatment with an indicator for high versus low perceived availability of the least preferred service.

Estimation   As the endline survey was administered to a subset of individuals characterized by a larger share of tenants than the overall sample, and there is selective attrition between the baseline and endline sample, the self-reported endline results may not be fully representative for the overall sample.[59] Taking this into consideration, I show two different specifications for each of the self-reported outcomes of interest. In each table, I first present the results of the respective regressions for the endline subsample without any adjustments (Column 1). These results can be cautionary interpreted as average effects for a randomly selected sample of owners and tenants. Second, I use inverse probability weights adjusting for the likelihood of being interviewed at endline conditional on ownership status and ward—the two main characteristics differing between the endline and the overall sample (Column 2). Assuming that these two characteristics predict selection into the endline sample, this should adjust the regressions to make them representative for the entire sample considered in the tax compliance analysis.[60] In each table, every row presents results from a separate regression and shows the effects of being treated on the respective index or index component.

Perceived Government Performance   I present the treatment effects on perceived government performance in Table 1.7. Column (1) shows that the effects of receiving information about public services on perceived performance of the local government are positive but insignificant, both when looking at the overall index, as well as at the five variables that feed into the index. Similarly, there are no effects of the treatment on perceived responsiveness of the government. These results are largely consistent with the specification using inverse probability weights in Column (2). In this specification, the effect on confidence in the Freetown City Council (FCC) is positive and significant, suggesting that specific, verifiable information about public services provided can indeed lead to increased trust in the local government. There are

---

[59] More information on the measurement challenges related to the endline sample can be found in 1.B.

[60] In addition, I use Horowitz-Manski bounds to look at the range of potential treatment effects in case the individuals who did not respond at endline are those with extreme outcomes. That is, for each individual who did not respond at endline, I replace the missing outcome value with the maximum value that this outcome takes for treated and the minimum for untreated individuals to calculate the upper bound of the treatment effect and vice versa for the lower bound. The results of these regressions are reported in Appendix Tables B12, B13, B14, B15, and B19. While these bounds can be useful to determine the range of potential effects, they are based on extreme cases with respect to selective attrition and are as such often very broad. Thus, they do not necessarily facilitate to draw conclusions about actual effect sizes.

no effects on perceived government performance among properties with above median value, indicating that this channel cannot explain the positive results on payments made for these properties.[61] While the signs on all coefficients suggest a weak positive link between the treatment and perceived government performance, a message mentioning a single example of public service provision does not seem to be enough to significantly shift attitudes towards the government overall.

Table 1.7: Treatment effects on perceived government performance

	Original	Inverse prob. weights
	(1)	(2)
Performance of local government index	0.051	0.051
	(0.036)	(0.038)
	[2,684]	[2,684]
- Confidence in ward councilor	0.032	0.024
	(0.036)	(0.039)
	[2,684]	[2,684]
- Confidence in mayor	0.025	0.019
	(0.037)	(0.040)
	[2,684]	[2,684]
- Confidence in FCC	0.050	0.066*
	(0.037)	(0.039)
	[2,684]	[2,684]
- Approval of ward councilor	0.050	0.045
	(0.036)	(0.039)
	[2,684]	[2,684]
- Approval of mayor	0.028	0.029
	(0.035)	(0.037)
	[2,684]	[2,684]
Local government responsiveness (std.)	0.043	0.036
	(0.033)	(0.035)
	[2,610]	[2,610]

Notes: Robust standard errors in parentheses. Number of observations in brackets. Significance levels are indicated by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Regressions include strata, enumerator, and ward fixed effects and control for respondent age, multiple property use, and the respective outcome variable at baseline. Appendix Table B12 presents results including Horowitz-Manski bounds.

---

[61] See Appendix Tables B13 and B14 for the effects on below and above median value properties.

Perceived Availability of and Satisfaction With Public Service Provision   Table 1.8 presents
the results of treatment effects on perceived availability of and satisfaction with local level pub-
lic service provision. The results show that, while the information treatment does not affect
perceived availability of services, it significantly increases satisfaction with the provision of
services of one's most and least preferred service, as well as service provision more generally.

Table 1.8: Treatment effects on perceived availability of and satisfaction with public services

	Original	Inverse prob. weights
	(1)	(2)
Availability of public services index	0.034	0.034
	(0.029)	(0.031)
	[2,684]	[2,684]
Availability of most preferred service (std.)	0.037	0.054
	(0.038)	(0.040)
	[2,677]	[2,677]
Availability of least preferred service (std.)	-0.005	-0.005
	(0.038)	(0.040)
	[2,665]	[2,665]
Satisfaction with public services index	0.058**	0.057*
	(0.029)	(0.031)
	[2,683]	[2,683]
Satisfaction with most preferred service (std.)	0.065*	0.090**
	(0.037)	(0.040)
	[2,676]	[2,676]
Satisfaction with least preferred service (std.)	0.071*	0.064*
	(0.037)	(0.039)
	[2,669]	[2,669]

Notes: Robust standard errors in parentheses. Number of observations in brackets. Significance levels are indi-
cated by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Regressions include strata, enumerator, and ward fixed effects
and control for respondent age, multiple property use, and the respective outcome variable at baseline. Appendix
Table B15 presents results including Horowitz-Manski bounds.

As expected given the content of the respective treatment messages, Table 1.9 shows a pos-
itive effect on satisfaction with the most preferred service in the match group, and a positive
effect on satisfaction with the least preferred service in the mismatch group. However, the co-
efficients are not significantly different across treatment arms. The treatment also significantly
increases satisfaction with service provision overall, measured by an index including all eight
types of services provided by the local government. This suggests that the treatment not only af-
fects satisfaction with the type of service one was informed about, but also generates spillovers

on satisfaction with other services—potentially due to increased awareness related to public service provision by the FCC post-treatment.

Table 1.9: Treatment effects on satisfaction with local level public service provision

	(1) Satisfaction with public services index	(2) Satisfaction with most preferred service (std.)	(3) Satisfaction with least preferred service (std.)
Panel A. Overall treatment effect			
Treated	0.066**	0.062*	0.078**
	(0.028)	(0.036)	(0.035)
Panel B. Treatment effect by arm			
Match	0.058*	0.080*	0.071*
	(0.033)	(0.042)	(0.041)
Mismatch	0.084**	0.059	0.086**
	(0.034)	(0.042)	(0.042)
p-val. Match = Mismatch	0.430	0.602	0.710
Obs.	2,664	2,660	2,659
Control mean	-0.000	-0.000	0.000
SD	1.000	1.000	1.000

Notes: Robust standard errors in parentheses. Significance levels are indicated by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Regressions include strata, enumerator, and ward fixed effects and control for respondent age, multiple property use, and the respective outcome variable at baseline. In addition, regressions control for public service preferences (most, least), linear distance to closest services (most, least), perceived availability of services at baseline (most, least), property value, and property ownership.

Interestingly, the effect of the mismatch treatment on satisfaction with one's least preferred service is entirely driven by residents of above median value properties.[62] Thus, the large and significant effect of the mismatch treatment on the amount of tax paid by residents of above median value properties could be driven by an increase in satisfaction with these services. This finding is in line with the hypothesis that individuals in the mismatch treatment are ex ante less informed about the treatment service, so that the treatment affects priors, and as such tax compliance, only within the mismatch group.

Finally, I also look at the role of perceived availability of services at baseline. Whereas perceived availability correlates with elicited preferences, it does not seem to affect the reaction to the treatment (see Appendix Table B18). This result should, however, be interpreted with caution, as comparably little variation in baseline perceived availability and individuals' generally low levels of perceived availability may prevent the identification of interaction effects along this dimension.

These results suggest that giving examples of one specific service provided by the local

---

[62] See Appendix Tables B16 and B17 for effects on satisfaction by property value.

government does not shift perceptions about the overall availability of this or any other type of service. However, it increases overall satisfaction with the provision of services by the local government. While part of this effect may be driven by new information about the existence of a service as argued above, the treatment may also increase the salience of who provides certain services. Attributability could thus play an important role in strengthening the fiscal social contract. The salience of politicians' performance with respect to public service provision can also be a crucial predictor of re-election chances.[63] The treatment is thus of direct interest to politicians who want to increase the visibility of their policy agenda.

1.9   Alternative Explanations

Tax Awareness   Given the nature of the intervention, the treatment could affect tax compliance through several other mechanisms than through the relevance of the information about public services itself. For instance, the intervention may simply raise awareness about an individual's duty to pay taxes and nudge individuals who forgot to do so or delayed payment into paying. Several pieces of evidence speak against this interpretation. First, the intervention mentioned neither the duty to pay property tax nor the payment deadline. The only sentence linked to taxes simply stated that "[t]he FCC has been raising necessary funds [to provide public services] through different sources, ranging from donor funds to property tax, local tax, and business license tax."[64]   If anything, this sentence underlines the necessity of raising tax revenue to provide public services and as such emphasizes the fiscal exchange channel.

Second, I show that the treatment effects are driven by individuals who care disproportionally about the provision of public services. In addition to official property taxes, individuals in Freetown often contribute to the provision of community services via informal contributions in cash, kind, or labor. These payments should not be overlooked, as they not only increase the tax burden of households, but also crucially contribute to the development of cities and countries (Olken and Singhal, 2011). At baseline, 85% of individuals reported that at least one of their household members made such a contribution in the previous year. Looking at the heterogeneity of effects by whether an informal tax payment was made or not reveals that the treatment effects are driven by individuals in whose household an informal contribution was made (see Appendix Table B20). These results suggest that individuals who value the provision of public services more are more likely to react to the treatment, underlining once again the importance

---

[63] For instance, Ajzenman and Durante (2023) shows that voters assigned to a ballot station in a school with poor infrastructure are less likely to vote for the incumbent politician than voters assigned to schools with higher quality infrastructure. The paper argues that this effect is driven by increased salience of the incumbent's poor performance with respect to the education sector in poor infrastructure schools.

[64] For individuals in the mismatch group, information about the provided service was further introduced by saying that "part of these funds have been used to improve" the respective service.

of the fiscal exchange channel as a driver of the observed effects. Rather than disrupting it, informal tax payments seem to reinforce the social contract with the state.

Finally, as has been illustrated at length in Sections 1.7.1 and 1.7.2, the content of the treatment message actually matters for whether the treatment does or does not affect tax compliance. If the treatment call simply made individuals aware of their tax duty, we should see a base treatment effect independent of the message's content. Instead I find that effects among lower-value properties are driven entirely by information aligning with perceived personal benefit. Whereas the fiscal exchange mechanism is somewhat less obvious for individuals residing in above median value properties, the fact that the match treatment does not affect tax compliance at all does speak against a pure awareness effect—which should be observed equally across treatment groups. It is thus much more likely that the treatment affected tax compliance through a change in perceived fiscal exchange induced by the information about public service provision than simply by bringing attention to the tax duty independently of the message's content.

Perceived Extractive Capacity   A second way of thinking about the treatment effect is through an effect on perceived extractive capacity of the government. Receiving personalized information about public services may induce people to (falsely) believe that the government now possesses more information about citizens that could facilitate the collection of tax revenue and the tracking of evaders. This makes evasion more costly and as such has the potential to increase tax compliance. Following Weigel (2020), I look at the effects of the intervention on *perceived extractive capacity* of the government using an index that reflects the perceived level of information the government has about taxpayers. his index increases in whether individuals believe that the Freetown City Council (FCC) knows the location of their house, whether their neighbors paid taxes, their occupation, and their income level.[65]

Table 1.10 illustrates the effects of being treated on the index (Row 1), as well as the standardized components of the index (Rows 2 - 5). Overall, the treatment does not affect perceived information on taxpayers. However, looking at the components separately shows that treated individuals are more likely to believe that the FCC knows whether their neighbors paid taxes as well as their occupation than individuals in the control group.

These results suggest that treated individuals may in fact, at least to a certain extent, believe the government more able to extract tax revenue than individuals in the control group. While I cannot entirely exclude that the treatment effect on tax compliance is (partly) driven by this increase, the effects on perceived FCC knowledge of one's occupation are not driven by the groups of individuals who increase compliance in response to the treatment. Residents of low-value properties who received information about the provision of access to water at below

---

[65]   As before, this index is constructed by standardizing each component of the index and taking the unweighted average of all components.

Table 1.10: Treatment effects on perceived government information on taxpayers

	Original	Inverse prob. weights
	(1)	(2)
Information on taxpayers index	0.035	0.043
	(0.037)	(0.040)
	[2,683]	[2,683]
HH location	-0.013	0.001
	(0.032)	(0.037)
	[2,683]	[2,683]
Compliance	0.048	0.052*
	(0.029)	(0.030)
	[2,683]	[2,683]
Occupation	0.077**	0.074**
	(0.035)	(0.037)
	[2,683]	[2,683]
Income	-0.038	-0.039
	(0.031)	(0.033)
	[2,683]	[2,683]

Notes: Robust standard errors in parentheses. Number of observations in brackets. Significance levels are indicated by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Regressions include strata, enumerator, and ward fixed effects and control for respondent age, multiple property use, and the respective outcome variable at baseline. Appendix Table B19 presents results including Horowitz-Manski bounds.

median distance (see Appendix Table B21) and residents of above median value properties, in particular those who received the mismatch treatment, do not react to the treatment by adjusting indicators of perceived extractive capacity (see Appendix Tables B22 and B23). This indicates that perceived extractive capacity does not drive the compliance results related to the information treatment and supports the conclusion that perceived fiscal exchange actually matters for tax compliance.

1.10 Additional Results

1.10.1 Owners Versus Tenants

Depending on who was reached at baseline, information about public services was provided either to the property owner—who is legally obliged to pay property taxes—or a resident without such obligations, such as an official tenant, a family member of the owner, or a caretaker. Given

that the link between the treatment and the payment of taxes is less strong if the latter is the case, one could expect the treatment effects to be driven entirely by property owners receiving information.

However, although tenants are not legally obliged to pay property tax, they are allowed do so on behalf of the owner once the payment deadline has passed. Payments made by the tenant can be deducted from the rent due.[66] More than that, the FCC has the right to seize the property, including the tenant's movable property, if no payment is made after receiving a warning letter. While this is unlikely to actually happen, it may give tenants reason to negotiate about tax payments with their landlord, especially upon receiving the public service treatment.

Table 1.11 shows that, on average, the treatment indeed affects tax compliance significantly only if the owner received information. However, looking at the two treatment arms separately shows significantly larger tax payments both among properties in which owners and tenants received information, specifically if they received information misaligned with respondent preferences. For owners, this effect may be driven by the information component of the mismatch treatment. As tenants can legally deduct property tax payments made in their landlord's stead from their rent only after the payment deadline, they are unlikely to make payments on their own prior to the deadline. Payments made for properties in which tenants received information are thus probably either made by the owner, or in accordance with the owner. If tenants communicate the provided information to their landlords and payments are made by the latter, the large reaction to the mismatch treatment could also be explained by differences in preferences between tenants and owners, with the latter responding to information about services aligned with their preferences.

1.10.2 Late Compliers Versus Non-Compliers

So far, this paper has shown that being treated increases tax compliance by the deadline. However, it is unclear whether these effects are driven by individuals who otherwise would have paid later (late compliers) or individuals who would not have paid at all (non-compliers). To shed light on this question, I look at the effects of information about public services on tax compliance throughout the 2022 tax cycle. Table 1.12 shows that the effects on the amount of tax paid persist over time. At the end of the year, individuals living in below median value properties who received information about their most preferred service pay on average 21% more taxes than residents of below median value properties in the control group. The effect is on average positive, but insignificant among treated individuals in the mismatch group. As before, the effects are largely driven by individuals who receive information about water at below median distance and decide to pay taxes in response to this information (see Appendix Figure B5).

---

[66] See https://fcc.gov.sl/property-reform-faq/

Table 1.11: Treatment effects on tax compliance by property ownership (September 2022)

	Amount paid (SLE)		Paid any tax (d)		Amount paid (SLE, cond.)	
	(1) Owner	(2) Tenant	(3) Owner	(4) Tenant	(5) Owner	(6) Tenant
Treated	16.717*	34.830	0.008	0.019	96.785**	159.050
	(9.873)	(25.258)	(0.009)	(0.017)	(45.513)	(188.848)
p-val T S1=S2	0.304		0.321		0.472	
Match	8.169	2.101	0.007	0.014	30.286	45.707
	(10.181)	(27.246)	(0.011)	(0.020)	(45.357)	(199.276)
Mismatch	23.212*	68.896*	0.010	0.030	149.210**	253.190
	(14.044)	(41.098)	(0.011)	(0.020)	(61.444)	(225.926)
p-val. Match = Mismatch	0.289	0.132	0.775	0.437	0.029	0.337
p-val. Match S1 = S2	0.725		0.689		0.102	
p-val. Mismatch S1 = S2	0.121		0.205		0.915	
Obs.	4,279	1,190	4,279	1,190	928	223
Control mean	115.167	131.709	0.214	0.155	537.448	848.794
SD	373.943	546.626	0.410	0.363	653.100	1154.713

Notes: Robust standard errors in parentheses. Significance levels are indicated by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Regressions include strata, enumerator, and ward fixed effects and control for respondent age, multiple property use, and the respective outcome variable at baseline. In addition, regressions control for public service preferences (most, least), linear distance to closest services (most, least), perceived availability of services at baseline (most, least), property value, and property ownership.

Table 1.12: Treatment effects on tax compliance by property value (December 2022)

	Amount paid (SLE)		Paid any tax (d)		Amount paid (SLE, cond.)	
	(1) < median	(2) > median	(3) < median	(4) > median	(5) < median	(6) > median
Treated	4.167**	27.533	0.011	0.002	9.641	52.642
	(1.917)	(28.944)	(0.011)	(0.014)	(7.705)	(93.100)
p-val T S1=S2	0.233		0.808		0.994	
Match	5.252**	-8.395	0.015	-0.002	12.367	-15.680
	(2.385)	(28.146)	(0.013)	(0.016)	(9.654)	(99.559)
Mismatch	3.306	73.971*	0.009	0.009	4.283	112.792
	(2.291)	(39.132)	(0.013)	(0.016)	(9.905)	(107.910)
p-val. Match = Mismatch	0.449	0.016	0.663	0.487	0.465	0.143
p-val. Match S1 = S2	0.692		0.994		0.673	
p-val. Mismatch S1 = S2	0.127		0.677		0.719	
Obs.	2,736	2,733	2,736	2,733	468	773
Control mean	25.084	258.728	0.163	0.273	153.596	946.319
SD	62.022	741.296	0.370	0.446	61.771	1167.218

Notes: Robust standard errors in parentheses. Significance levels are indicated by $*$ $p < 0.10$, $**$ $p < 0.05$, $***$ $p < 0.01$. Regressions include strata, enumerator, and ward fixed effects and control for respondent age, multiple property use, and the respective outcome variable at baseline. In addition, regressions control for public service preferences (most, least), perceived availability of services at baseline (most, least), linear distance to closest services (most, least), property value, and property ownership.

This suggests that the treatment does not only induce these individuals to pay taxes earlier, but actually motivates taxpayers who otherwise would not pay taxes to comply.

Similarly, the positive effect of the mismatch treatment among residents of above median value properties persists over time. By the end of the 2022 tax cycle, these individuals pay on average 29% more taxes—an effect that has substantial influence on overall tax revenue. The increase in payments seems to be driven by effects on both the extensive and intensive margin. However, the effects cannot be identified separately at conventional levels of significance. Overall, this suggests that individuals residing in above median value properties who receive the mismatch treatment pay significantly more taxes than individuals in the control group instead of simply paying (more) earlier.

Finally, it is also worth looking at whether the treatment effect persists beyond the tax year in which the intervention took place. Looking at tax payments for the 2023 tax cycle until June 20, 2023 shows that the treatment does not affect tax payments more than a year after the treatment (see Appendix Table B24). Taken together, these results suggest that light-touch interventions like the one implemented in this paper have important effects on tax compliance in the short- and medium run. However, a one time intervention is not enough to change individual tax compliance behavior in the long run. Whether and how repeated information treatments could shape tax compliance in the long run deserves further investigation.

1.11   Conclusion

This paper examines the effects of specific, targeted information about public service provision on tax compliance. I find that treated individuals pay on average 20% more taxes on time than individuals in the control group. The effect is driven by increases on both the extensive and intensive margin of tax compliance. Individuals residing in low-value properties, who depend heavily on public service provision, react to information that signals personal benefit. in particular, these individuals increase their likelihood of paying taxes on time by 7–16 percentage points when receiving information about an easily accessible service with unambiguous value for all citizens, namely water. The effect of public service information persists throughout the tax cycle, suggesting that messages emphasizing fiscal exchange have the potential to not only shift the timing of tax payments, but actually increase compliance.

These results emphasize the importance of the underlying level of public service supply, illustrating that interventions showcasing government effort to provide services do not increase tax compliance among those who most depend on public services when services are too far away. Hence, increasing the density of public services, and as such service provision more generally, is a crucial condition for fiscal exchange messages to affect tax compliance among

those most in need. By itself, increasing public service provision has been shown to positively affect tax compliance. Informing individuals about the latter can have complementary effects, thus increasing the returns to the provision of public services for local governments.

Residents of high-value properties, on the other hand, increase tax compliance in response to information about their least preferred service. These effects are not driven by the specific content of the message, suggesting that benefiting personally from the service one is informed about may be less relevant for residents of properties with better average access to public services. However, individuals in the mismatch group update their priors about local level public service provision and are ex-post more satisfied with the provision of their least preferred type of service. The effect of the mismatch treatment can thus be interpreted as reaction to a positive signal of government performance with respect to public service provision that is independent of personal benefit.

My results also show that individuals residing in high-value properties drive the effects on overall tax revenue. This is an important finding, suggesting that policy makers with the goal of increasing tax revenue should not only focus on providing information about public services to those most in need of the latter. By contrast, providing information that signals the government's capacity to deliver public services to richer individuals has the potential to generate substantially higher overall returns for the government.

Bibliography

Abbiati, L., Antinyan, A., and Corazzini, L. (2020). A survey experiment on information, taxpayer preferences, and perceived adequacy of the tax burden. *Heliyon*, 6(3).

Ajzenman, N. and Durante, R. (2023). Salience and accountability: School infrastructure and last-minute electoral punishment. *The Economic Journal*, 133(649):460–476.

Alm, J., Jackson, B. R., and McKee, M. (1993). Fiscal exchange, collective decision institutions, and tax compliance. *Journal of Economic Behavior & Organization*, 22(3):285–303.

Antinyan, A. and Asatryan, Z. (2020). Nudging for tax compliance: A meta-analysis. CESifo Working Paper No. 8500.

Balán, P., Bergeron, A., Tourek, G., and Weigel, J. L. (2022). Local elites as state capacity: How city chiefs use local information to increase tax compliance in the Democratic Republic of the Congo. *American Economic Review*, 112(3):762–97.

Bergeron, A., Tourek, G. Z., and Weigel, J. L. (2023). The state capacity ceiling on tax rates: Evidence from randomized tax abatements in the DRC. NBER Working Paper No. w31685.

Bergolo, M., Ceni, R., Cruces, G., Giaccobasso, M., and Perez-Truglia, R. (2023). Tax audits as scarecrows: Evidence from a large-scale field experiment. *American Economic Journal: Economic Policy*, 15(1):110–53.

Besley, T. (2020). State capacity, reciprocity, and the social contract. *Econometrica*, 88(4):1307–1335.

Besley, T. and Persson, T. (2009). The origins of state capacity: Property rights, taxation, and politics. *American Economic Review*, 99(4):1218–44.

Besley, T. and Persson, T. (2013). Taxation and development. In *Handbook of Public Economics*, volume 5, pages 51–110, Amsterdam. Elsevier.

Blumenthal, M., Christian, C., and Slemrod, J. (2001). Do normative appeals affect tax compliance? Evidence from a controlled experiment in Minnesota. *National Tax Journal*, 54(1):125–138.

Bodea, C. and LeBas, A. (2016). The origins of voluntary compliance: Attitudes toward taxation in urban Nigeria. *British Journal of Political Science*, 46(1):215–238.

Bott, K. M., Cappelen, A. W., Sørensen, E. Ø., and Tungodden, B. (2020). You've got mail: A randomized field experiment on tax evasion. *Management Science*, 66(7):2801–2819.

Brockmeyer, A., Estefan, A., Arras, K. R., and Serrato, J. C. S. (2023). Taxing property in developing countries: Theory and evidence from Mexico. NBER Working Paper No. w28637.

Casal, S., Kogler, C., Mittone, L., and Kirchler, E. (2016). Tax compliance depends on voice of taxpayers. *Journal of Economic Psychology*, 56:141–150.

Castro, L. and Scartascini, C. (2015). Tax compliance and enforcement in the pampas evidence from a field experiment. *Journal of Economic Behavior & Organization*, 116:65–82.

Cohen, I. (2023). Technology and the state: Building capacity to tax via text. SSRN Working Paper No. 4523287.

Collin, M. E., Di Maro, V., Evans, D. K., and Manang, F. (2022). Property tax compliance in Tanzania: Can nudges help? CGD Working Paper 621.

De Neve, J.-E., Imbert, C., Spinnewijn, J., Tsankova, T., and Luts, M. (2021). How to improve tax compliance? Evidence from population-wide experiments in Belgium. *Journal of Political Economy*, 129(5):1425–1463.

Dom, R., Custers, A., Davenport, S., and Prichard, W. (2022). Taxing at the local level. In *Innovations in Tax Compliance: Building Trust, Navigating Politics, and Tailoring Reform*, chapter 6, pages 141–170. World Bank Publications.

Fernandez, M., Gonzalez-Navarro, M., and Quintana-Domeque, C. (2023). Local public goods and property tax compliance: Evidence from residential street pavement. Unpublished Manuscript, presented at the EEA-ESEM congress 2023.

Fish, P. (2018). Practical guidance note: Training manual for implementing property tax reform with a points-based valuation. ICTD African Tax Administration Paper 2.

Freetown City Council (2019). *Transform Freetown: An Overview | 2019 – 2022*. FCC, Freetown.

Giaccobasso, M., Nathan, B. C., Perez-Truglia, R., and Zentner, A. (2023). Where do my tax dollars go? Tax morale effects of perceived government spending. NBER Working Paper No. w29789.

Grieco, K., Bakarr Kamara, A., Meriggi, N. F., Michel, J., Prichard, W., and Stewart-Wilson, G. (2019). Simplifying property tax administration in africa: Piloting a points-based valuation in Freetown, Sierra Leone. IGC Policy Brief 39427.

Hallsworth, M., List, J. A., Metcalfe, R. D., and Vlaev, I. (2017). The behavioralist as tax collector: Using natural field experiments to enhance tax compliance. *Journal of Public Economics*, 148:14–31.

Hanousek, J. and Palda, F. (2004). Quality of government services and the civic duty to pay taxes in the Czech and Slovak republics, and other transition countries. *Kyklos*, 57(2):237–252.

Hernandez, M., Jamison, J., Korczyc, E., Mazar, N., and Sormani, R. (2017). *Applying behavioral insights to improve tax collection: Experimental Evidence from Poland*. World Bank, Washington, DC.

Hobbes, T. (1651). *Leviathan or the Matter, Forme, & Power of a Common-wealth Ecclesiasticall and Civill*. Andrew Crooke, London.

Hofmann, E., Hoelzl, E., and Kirchler, E. (2008). Preconditions of voluntary tax compliance: Knowledge and evaluation of taxation, norms, fairness, and motivation to cooperate. *Zeitschrift für Psychologie/Journal of Psychology*, 216(4):209–217.

Jibao, S. S. (2017). Sierra Leone. In Franzsen, R. and McCluskey, M., editors, *Property tax in Africa: Status, challenges, and prospects*, pages 264–376. Lincoln Institute of Land Policy, Cambridge, Massachusetts.

Kaldor, N. (1963). Taxation for economic development. *The Journal of Modern African Studies*, 1(1):7–23.

Kamara, A. B., Meriggi, N., and Prichard, W. (2020). Freetown just implemented a new property tax system that could quintuple revenue. International Centre for Tax and Development Blog [May 22, 2020].

Khan, A., Khawja, A., Olken, B., and Shaukat, M. (2022). Rebuilding the social compact: Urban service delivery and property taxes in Pakistan. IGC Final Report PAK-22003.

Khan, A. Q., Khwaja, A. I., and Olken, B. A. (2019). Making moves matter: Experimental evidence on incentivizing bureaucrats through performance-based postings. *American Economic Review*, 109(1):237–70.

Knebelmann, J. and Pouliquen, V. (2023). Discretion versus algorithms: Bureaucrats and tax equity in Senegal. Unpublished Manuscript, Massachusetts Institute of Technology.

Krause, B. (2020). Balancing purse and peace: Tax collection, public goods and protests. Unpublished Manuscript, University of California Berkley.

Lamberton, C., De Neve, J.-E., and Norton, M. I. (2018). The power of voice in stimulating morality: Eliciting taxpayer preferences increases tax compliance. *Journal of Consumer Psychology*, 28(2):310–328.

Levi, M. (1988). *Of Rule and Revenue*. University of California Press, Berkeley, CA.

Locke, J. (1690). *Two Treatises of Government: In the Former, The False Principles and Foundation of Sir Robert Filmer, And His Followers, are Detected and Overthrown. The Latter is an Essay concerning The True Original, Extent, and End of Civil-Government*. Awnsham Churchill, London.

Luttmer, E. F. and Singhal, M. (2014). Tax morale. *Journal of Economic Perspectives*, 28(4):149–68.

Mascagni, G. (2018). From the lab to the field: A review of tax experiments. *Journal of Economic Surveys*, 32(2):273–301.

Mascagni, G. and Nell, C. (2022). Tax compliance in Rwanda: Evidence from a message field experiment. *Economic Development and Cultural Change*, 70(2):587–623.

Moore, M., Prichard, W., and Fjeldstad, O.-H. (2018). *Taxing Africa: Coercion, reform and development*. Zed Books Ltd, London.

OECD (2019). *Tax Morale: What Drives People And Businesses To Pay Tax?* OECD Publishing, Paris.

OECD (2022). Revenue statistics in Africa 2022 – Sierra Leone. Available online [last accessed April 23, 2024].

Olken, B. A. and Singhal, M. (2011). Informal taxation. *American Economic Journal: Applied Economics*, 3(4):1–28.

Prichard, W. (2017). Linking property tax revenue and public services. ICTD Summary Brief 13.

Rousseau, J. J. (1762). *Du contrat social; ou, Principes du droit politique*. Marc-Michel Rey, Amsterdam.

Sjoberg, F. M., Mellon, J., Peixoto, T. C., Hemker, J. Z., and Tsai, L. L. (2019). Voice and punishment: A global survey experiment on tax morale. World Bank Policy Research Working Paper No. 8855.

Slemrod, J. (2007). Cheating ourselves: The economics of tax evasion. *Journal of Economic perspectives*, 21(1):25–48.

United Nations, Department of Economic and Social Affairs, Population Division (2019). *World Urbanization Prospects: The 2018 Revision (ST/ESA/SER.A/420)*. United Nations, New York.

Van den Boogaard, V., Prichard, W., Beach, R., and Mohiuddin, F. (2020). Strengthening tax-accountability links: Fiscal transparency and taxpayer engagement in Ghana and Sierra Leone. ICTD Working Paper 114.

Weigel, J. L. (2020). The participation dividend of taxation: How citizens in Congo engage more with the state when it tries to tax them. *The Quarterly Journal of Economics*, 135(4):1849–1903.

Appendices

1.A Study Background and Data

Figure A1 illustrates total municipality level revenue, as well as property tax revenue (as a share of total revenue) in Freetown between 2019 and 2022.

Figure A1: Property tax revenue as compared to overall FCC budget
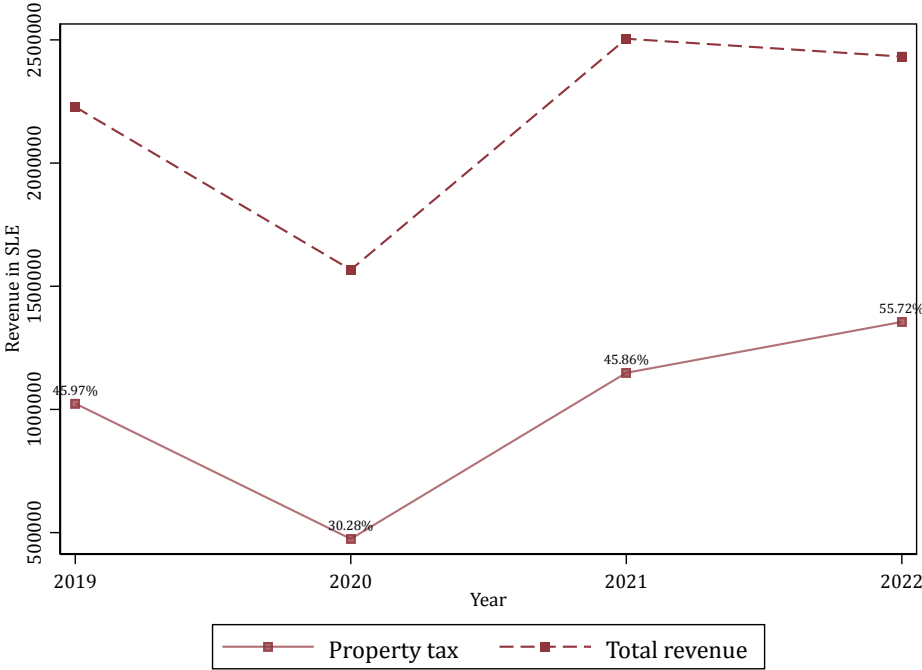
The following paragraphs explain the sampling and data collection procedures in more detail. Figure A3 at the end of the section illustrates the process graphically.

Sampling The target population of this project are owners and residents of registered domestic properties across Freetown. To sample a subset of these individuals, I use the database of registered properties generated during the recent implementation of the property tax reform. In January 2022, the database contained information about 103,407 registered properties, out

of which 97,846 are domestic properties.[67] The database contains information about the property value, annual and outstanding tax liabilities, and the exact, geo-referenced location of the property, as well as phone numbers of at least one owner or resident for 79,860 properties. Restricting the database to domestic properties with at least one phone number available, I am left with 75,687 properties.

To test the effect of information about public services among individuals for whom public service provision is considered most relevant, I restrict the sample to properties which are owned by an individual who does not own more than one property. Ownership of more than one registered property is used as an indicator for disproportionally wealthy individuals within the sample. Since wealthier citizens in Freetown tend to rely less on public service provision, and instead often make use of private providers or provide services for themselves, information on public service delivery is likely to be less relevant for these individuals. I proxy ownership of multiple properties using overlaps in owner names and phone numbers across properties.[68] Excluding properties owned by a person who possesses several properties, I am left with 46,844 properties.

To have sufficient power to test all outcomes of interest, I aimed at a final sample of 7,000 individuals. Given projections for response rates based on similar projects in the same setting, I randomly sampled 15,000 properties from the above-described final database as the respondent pool for this project.

Baseline   During the baseline data collection, enumerators conducted phone interviews with a total of 5,827 individuals.[69] The collection of baseline data had three main goals: (1) to collect individual level characteristics used for a stratified randomization of the treatment assignment, (2) to elicit public service preferences in order to provide targeted information during the intervention, and (3) to collect information on baseline attitudes towards public service provision, taxation, and the local government more broadly.

Due to budgetary constraints, the full baseline survey could be conducted only with a subset of individuals. The baseline data collection was therefore done in two parts. In the first part (batch 1), enumerators interviewed individuals who either own or reside in a sampled property. These respondents received the full length survey, including a variety of questions about the local government and their interaction with the latter. It was planned that these individuals

---

[67] This includes properties used for both domestic and commercial purposes as well as purely domestic properties.

[68] Given that names were recorded manually by different enumerators when assessing the properties, names are likely spelled differently across properties even if they belong to the same individual. I therefore generate a variable quantifying the extent to which two names with the same phone number overlap using continuous values from 0 to 1. If this variable is larger than or equal to 0.7, I consider two observations with the same phone number to belong to the same individual.

[69] Due to uncertainties concerning the funding at this stage, it was not feasible to continue the data collection until the envisioned 7,000 individuals were reached.

would receive an endline survey that would allow to assess the effects of the treatment on the respective variables. In total, 3,581 individuals were interviewed during this stage.

In the second part (batch 2), respondents received a shorter version of the baseline survey, which included questions about individual level characteristics, baseline public service preferences, and a small subset of relevant baseline attitudes, e.g., perceived availability of public services. For these individuals, focus was put on assessing the effects of the treatment on tax compliance using administrative data, and no endline survey was planned. Since effects on tax compliance (rather than attitudes) were expected to arise predominantly through an interaction with property owners—i.e., those legally obliged to pay property tax—the second part of the baseline data collection was restricted to property owners only. A total of 2,246 individuals were interviewed during this stage.

Given the design of the baseline data collection, property owners are oversampled with respect to the natural rate of reaching an owner rather than a tenant when calling a contact number provided to the Freetown City Council. Whereas 65% of respondents in the first batch are property owners, the final sample of individuals interviewed at baseline consists of 78% property owners.

I link the baseline data to the administrative tax data including property locations, as well as the dataset of pubic service locations. Prior to randomization, I exclude the 3% of individuals with the largest distance to their most preferred service and the 3% of individuals with the largest distance to their least preferred service. Similarly, individuals with missing data on their most or least preferred public service are excluded. This leaves me with a final sample of 5,494 individuals, 3,389 in batch 1 and 2,105 in batch 2.

Endline The endline survey was administered roughly 8 to 12 days after the intervention. Figure A2 illustrates the exact distribution of days between intervention and endline survey in the treatment group. Individuals in the control were interviewed parallel to those in the treatment group, with the aim of keeping the share of treated and control group individuals interviewed constant over the course of the data collection. The endline survey predominantly contained questions about attitudes towards the local government, taxation, and public service provision. In total, 2,684 individuals were interviewed. This represents 79% of individuals in batch 1.

Figure A2: Time span between intervention and endline (treatment group, days)
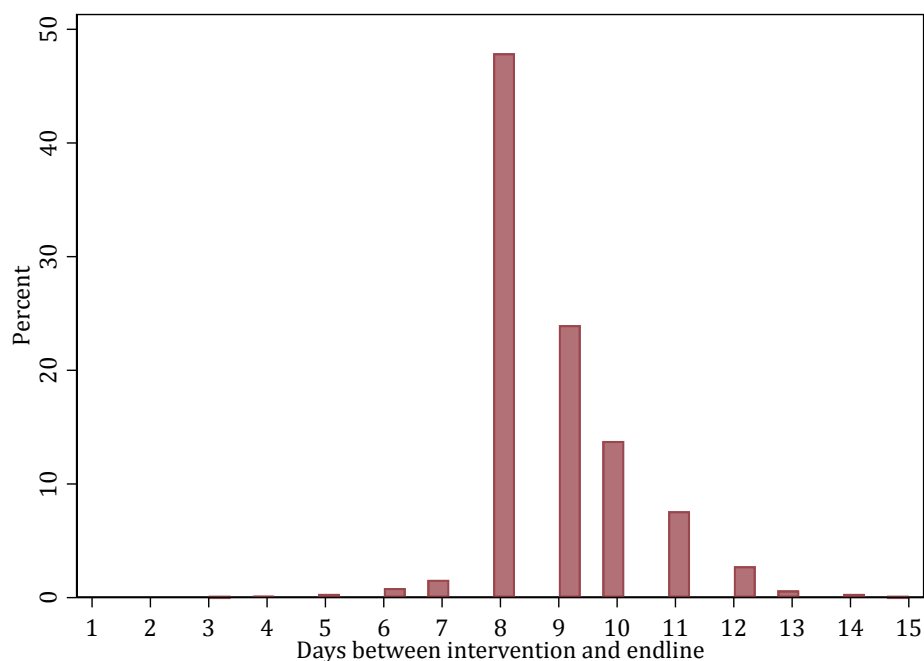
Notes: The figure illustrates the amount of time in days between the intervention and the endline survey for individuals who both received treatment and participated in the endline survey.

Table A1: Characteristics of eligible as compared to sampled properties

Variable	N	(1) Eligible Mean/SE	N	(2) Sampled Mean/SE	t-test Difference (1)-(2)
Assessed property value	31845	13818.404 [107.970]	14999	14070.650 [157.484]	-252.246
Property is both domestic and commercial/institutional	31845	0.050 [0.001]	14999	0.055 [0.002]	-0.005**
Ward group 1	31844	0.184 [0.002]	14999	0.193 [0.003]	-0.009**
Ward group 2	31844	0.135 [0.002]	14999	0.139 [0.003]	-0.004
Ward group 3	31844	0.133 [0.002]	14999	0.130 [0.003]	0.003
Ward group 4	31844	0.134 [0.002]	14999	0.128 [0.003]	0.005
Ward group 5	31844	0.192 [0.002]	14999	0.192 [0.003]	-0.000
Ward group 6	31844	0.222 [0.002]	14999	0.217 [0.003]	0.005

Notes: The value displayed for t-tests are the differences in the means across groups. Significance levels are indicated by $*$ $p < 0.10$, $**$ $p < 0.05$, $***$ $p < 0.01$. Ward groups refer to groupings of neighboring wards.

Figure A3: Research Design and Data Collection Diagram

46,844 eligible domestic properties
owned by single-property owners with a.l. 1 contact number available

31,844 properties
eligible, but not sampled

15,000 properties
randomly sampled for potential interviews with residents

9,173 properties
no interview conducted

5,827 baseline surveys
with owners and tenants

333 respondents
excluded before rand.

5,494 respondents
randomized into treatment and control

Long baseline survey
3,389 owners and tenants

Short baseline survey
2,105 owners

T_M
N=1,126

T_L
N=1,124

C
N=1,139

T_M
N=703

T_L
N=707

C
N=695

Attrition
N=705

Endline survey
N=2,684

Post-treatment administrative data
5,494 properties within the study sample

Administrative data
9,448 sampled properties
31,645 unsampled prop.

Notes: This figure illustrates the data collection procedure from sampling, over baseline data collection, randomization, and endline data collection, including information on the respective group sizes.

Figure A4: Regressions of service priorities on perceived distance to the closest service of each type and the amount of services provided within a 3km radius
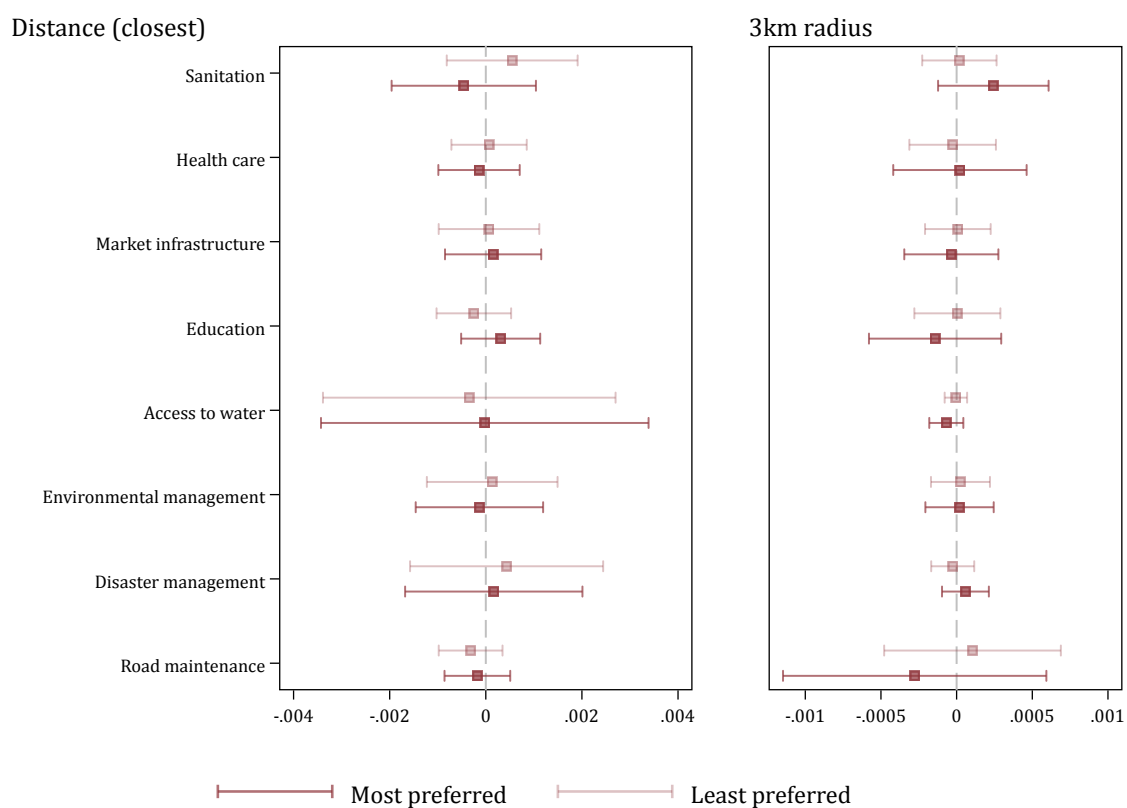
Notes: The figure shows correlations between preferences for a specific type of service and distance to the closest service of a that type recently provided by the FCC (left), as well as the amount of services of that type recently provided (right). Coefficients stem from simultaneous regressions and are shown with 95% confidence intervals. All regressions include baseline surveyor fixed effects.

Figure A5: Regressions of service priorities on perceived availability of services, satisfaction with services, property value, and property ownership
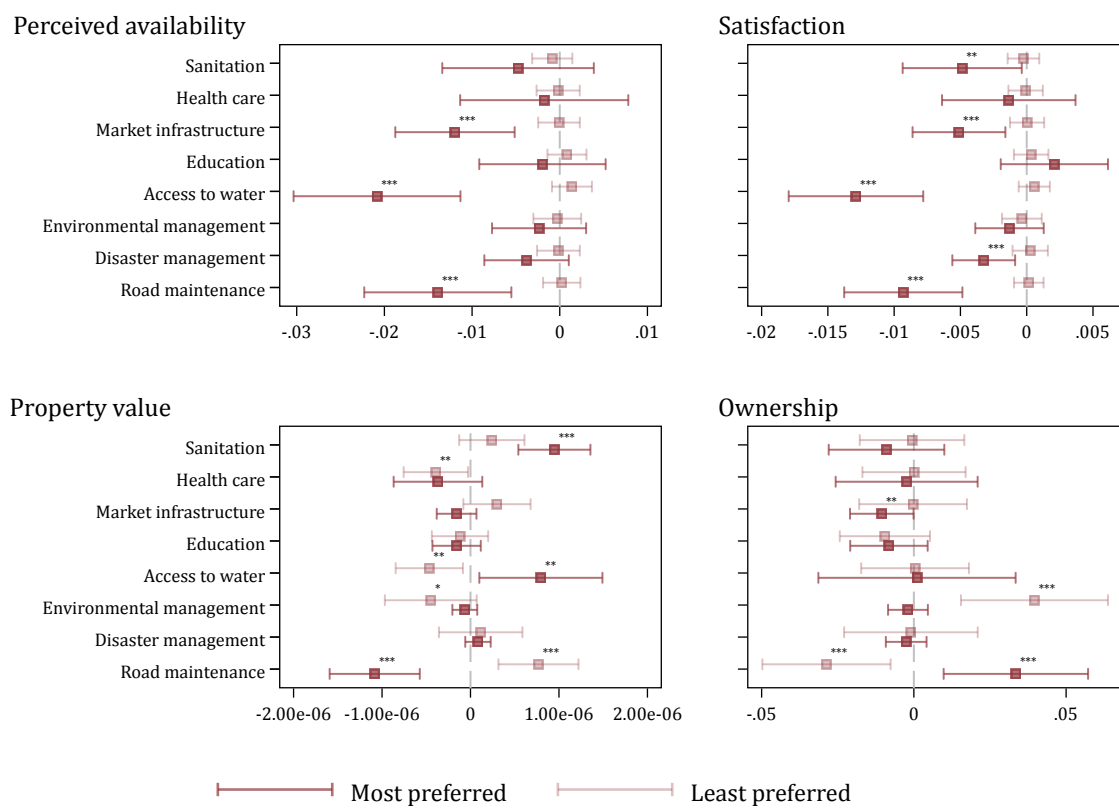
Notes: The figure shows correlations between preferences for a specific type of service and perceived availability of that type of service (upper left), satisfaction with that type of service (upper right), the value of the property in which respondents live (lower left), and ownership of the respective property (lower right). Coefficients stem from simultaneous regressions and are shown with 95% confidence intervals. All regressions include baseline surveyor fixed effects.

Figure A6: Intervention phone call script (SurveyCTO)

Today, we are calling you to provide some information about the progress of the Transform Freetown Agenda. In line with this agenda, the Freetown City Council works on improving access to public services across Freetown. The FCC has been raising the necessary funds through different sources, ranging from donor funds to property tax, local tax, and business license tax.

When we last called you, you said that **ACCESS TO WATER** was/were your top priority.

Today, we are calling you to provide some information about the progress of the Transform Freetown Agenda. In line with this agenda, the Freetown City Council works on improving access to public services across Freetown. The FCC has been raising the necessary funds through different sources, ranging from donor funds to property tax, local tax, and business license tax.

Part of these funds have been used to improve **ACCESS TO WATER** in several areas in Freetown.

ACCESS TO WATER has/have been improved, for example, at **Bottom Oku**, **around a 5-10 minutes walk** from your house. Improvements in/of **ACCESS TO WATER** typically include the installation of boreholes, water pumps, or milla tanks.

Do you know about the FCC's improvements in terms of **ACCESS TO WATER** done in this area?

○ Yes

○ No

Thank you very much for your time. Following this call, I will send you a short message on your phone to remind you of where ACCESS TO WATER has/have been improved by the FCC in your area. As I said, these improvements have been done at **Bottom Oku**, **around a 5-10 minutes walk** from your house.

Should you be interested into receiving more details, do not hesitate to get in touch with the FCC representative in your ward.

Thanks again for listening to me and for answering my questions.

(a) Part I - Match   (b) Part I - Mismatch   (c) Part II   (d) Part III

Notes: The figure shows the script used for the intervention call. Figures A6a and A6b illustrate the call's introduction for the match and mismatch treatment, respectively. Figure A6c shows the main part of the message. Figure A6d illustrates the final part of the call.

Figure A7: Follow-up short message (SMS)

TO
phone number

MESSAGE
Dear respondent name , thank you for your time! Let me remind you that the FCC has improved ACCESS TO WATER at Calaba Town Market, around a 10-20 minutes walk from your house.

Send SMS
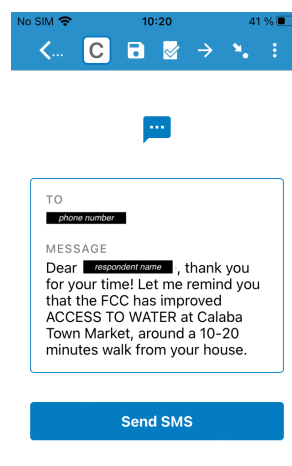
Notes: The figure shows an example of an SMS sent to treated respondents after the treatment call.

Figure A8: Type of service individuals where informed about by treatment arm
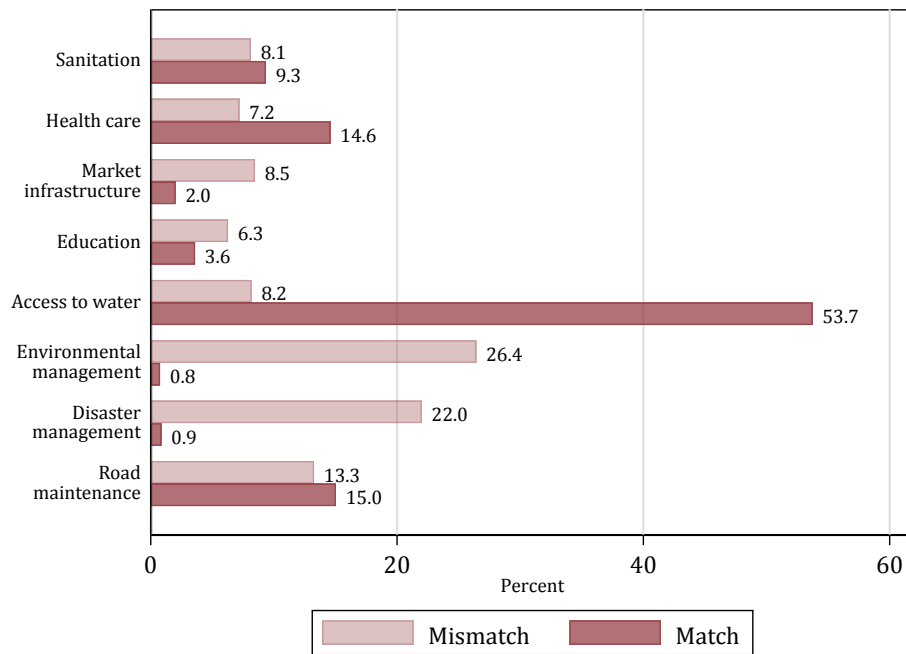
Notes: The figure shows the share of individuals in the match and mismatch group informed about a specific type of service.

Figure A9: Type of service individuals where informed about
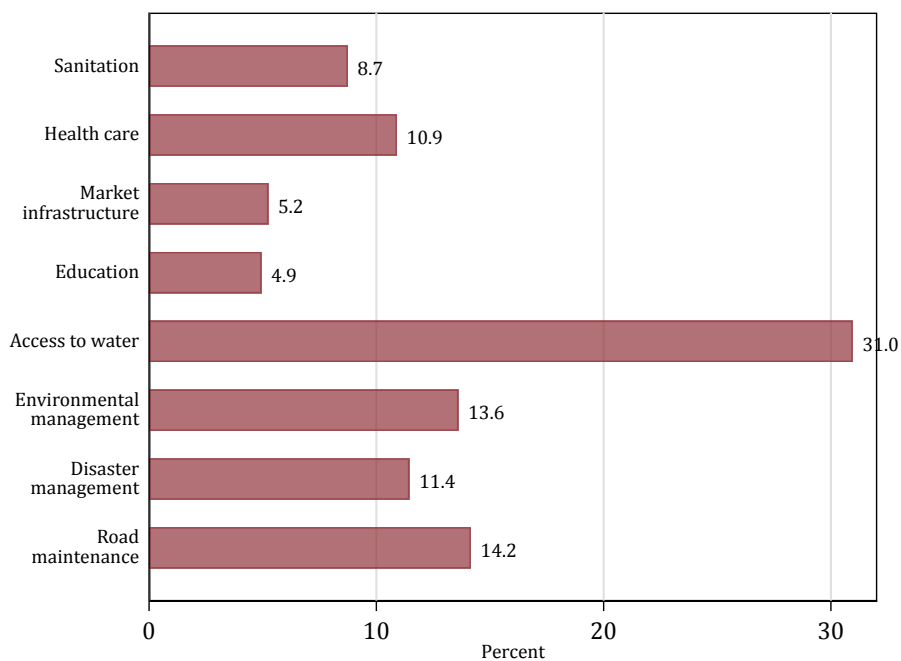
Notes: The figure shows the share of treated individuals informed about a specific type of service.

Table A2: Characteristics of eligible as compared to interviewed properties

Variable	N	(1) Eligible Mean/SE	N	(2) Interviewed Mean/SE	t-test Difference (1)-(2)
Assessed property value	41350	13914.932 [95.247]	5494	13780.543 [250.256]	134.389
Property is both domestic and commercial/institutional	41350	0.050 [0.001]	5494	0.058 [0.003]	-0.007**
Ward group 1	41349	0.192 [0.002]	5494	0.153 [0.005]	0.038***
Ward group 2	41349	0.134 [0.002]	5494	0.159 [0.005]	-0.025***
Ward group 3	41349	0.130 [0.002]	5494	0.145 [0.005]	-0.015***
Ward group 4	41349	0.133 [0.002]	5494	0.122 [0.004]	0.012**
Ward group 5	41349	0.190 [0.002]	5494	0.206 [0.005]	-0.015***
Ward group 6	41349	0.221 [0.002]	5494	0.215 [0.006]	0.006

Notes: The value displayed for t-tests are the differences in the means across groups. Significance levels are indicated by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Ward groups refer to groupings of neighboring wards.

Table A3: Baseline characteristics of individuals interviewed at endline compared to those not interviewed at endline

Variable	N	(1) No Mean/SE	N	(2) Yes Mean/SE	T-test Difference (1)-(2)
Female	2810	0.356 (0.009)	2684	0.349 (0.009)	0.006
Age	2804	5.813 (0.050)	2673	5.447 (0.050)	0.366***
Education level	2785	4.269 (0.044)	2682	4.208 (0.045)	0.061
In relationship	2806	0.870 (0.006)	2683	0.873 (0.006)	-0.003
Owner	2810	0.913 (0.005)	2684	0.644 (0.009)	0.269***
Monthly income	2399	3.351 (0.037)	2577	3.423 (0.036)	-0.071
Ward group 1	2810	0.148 (0.007)	2684	0.159 (0.007)	-0.011
Ward group 2	2810	0.150 (0.007)	2684	0.168 (0.007)	-0.018*
Ward group 3	2810	0.143 (0.007)	2684	0.147 (0.007)	-0.003
Ward group 4	2810	0.123 (0.006)	2684	0.120 (0.006)	0.002
Ward group 5	2810	0.214 (0.008)	2684	0.197 (0.008)	0.016
Ward group 6	2810	0.222 (0.008)	2684	0.208 (0.008)	0.013
Distance to most preferred service (km)	2810	1.476 (0.033)	2684	1.573 (0.035)	-0.098**
Distance to least preferred service (km)	2810	1.632 (0.028)	2684	1.680 (0.030)	-0.048
Most preferred service: water	2810	0.542 (0.009)	2684	0.509 (0.010)	0.033**
Least preferred service: water	2810	0.085 (0.005)	2684	0.082 (0.005)	0.003
Property value (SLL)	2810	1.45e+07 (3.64e+05)	2684	1.31e+07 (3.42e+05)	1.38e+06***

Notes: The value displayed for t-tests are the differences in the means across the groups. ***, **, and * indicate significance at the 1, 5, and 10 percent critical level.

Table A4: Balance of baseline characteristics across treatment groups

Variable	(1) Match N	(1) Match Mean/SE	(2) Mismatch N	(2) Mismatch Mean/SE	(3) Control N	(3) Control Mean/SE	T-test Difference (1)-(2)	T-test Difference (1)-(3)	T-test Difference (2)-(3)
Female	1829	0.353 (0.011)	1831	0.353 (0.011)	1834	0.352 (0.011)	-0.000	0.000	0.001
Age	1822	5.581 (0.061)	1827	5.732 (0.063)	1828	5.590 (0.060)	-0.152*	-0.010	0.142
Education level	1817	4.232 (0.055)	1821	4.273 (0.054)	1829	4.212 (0.054)	-0.041	0.021	0.061
In relationship	1826	0.867 (0.008)	1830	0.866 (0.008)	1833	0.883 (0.008)	0.001	-0.016	-0.017
Owner	1829	0.781 (0.010)	1831	0.785 (0.010)	1834	0.779 (0.010)	-0.004	0.003	0.006
Monthly income	1660	3.400 (0.045)	1659	3.374 (0.046)	1657	3.391 (0.044)	0.026	0.009	-0.017
Ward group 1	1829	0.153 (0.008)	1831	0.152 (0.008)	1834	0.155 (0.008)	0.000	-0.003	-0.003
Ward group 2	1829	0.158 (0.009)	1831	0.161 (0.009)	1834	0.158 (0.009)	-0.003	0.000	0.004
Ward group 3	1829	0.146 (0.008)	1831	0.147 (0.008)	1834	0.142 (0.008)	-0.001	0.004	0.006
Ward group 4	1829	0.121 (0.008)	1831	0.122 (0.008)	1834	0.122 (0.008)	-0.001	-0.001	-0.000
Ward group 5	1829	0.205 (0.009)	1831	0.205 (0.009)	1834	0.207 (0.009)	-0.000	-0.002	-0.002
Ward group 6	1829	0.218 (0.010)	1831	0.212 (0.010)	1834	0.216 (0.010)	0.006	0.002	-0.004
Distance to most preferred service (km)	1829	1.518 (0.042)	1831	1.537 (0.041)	1834	1.515 (0.041)	-0.019	0.003	0.022
Distance to least preferred service (km)	1829	1.644 (0.035)	1831	1.681 (0.036)	1834	1.643 (0.035)	-0.038	0.001	0.039
Most preferred service: water	1829	0.537 (0.012)	1831	0.522 (0.012)	1834	0.519 (0.012)	0.015	0.019	0.004
Least preferred service: water	1829	0.077 (0.006)	1831	0.082 (0.006)	1834	0.092 (0.007)	-0.005	-0.015	-0.010
Property value (SLL)	1829	1.40e+07 (4.55e+05)	1831	1.34e+07 (4.17e+05)	1834	1.40e+07 (4.27e+05)	5.93e+05	-1.45e+04	-6.08e+05

Notes: The value displayed for t-tests are the differences in the means across the groups. ***, **, and * indicate significance at the 1, 5, and 10 percent critical level.

Table A5: Balance of baseline characteristics across treatment arms among individuals who participated in the endline survey

Variable	(1) Match N	Mean/SE	(2) Mismatch N	Mean/SE	(3) Control N	Mean/SE	T-test Difference (1)-(2)	(1)-(3)	(2)-(3)
Female	900	0.342 (0.016)	883	0.350 (0.016)	901	0.356 (0.016)	-0.008	-0.014	-0.006
Age	898	5.362 (0.085)	880	5.619 (0.091)	895	5.363 (0.084)	-0.257**	-0.001	0.256**
Education level	898	4.259 (0.078)	883	4.227 (0.078)	901	4.138 (0.077)	0.033	0.122	0.089
In relationship	899	0.867 (0.011)	883	0.866 (0.011)	901	0.887 (0.011)	0.000	-0.020	-0.020
Owner	900	0.652 (0.016)	883	0.644 (0.016)	901	0.636 (0.016)	0.008	0.016	0.008
Monthly income	868	3.472 (0.062)	848	3.439 (0.063)	861	3.357 (0.062)	0.034	0.116	0.082
Ward group 1	900	0.151 (0.012)	883	0.153 (0.012)	901	0.173 (0.013)	-0.002	-0.022	-0.020
Ward group 2	900	0.160 (0.012)	883	0.176 (0.013)	901	0.169 (0.012)	-0.016	-0.009	0.007
Ward group 3	900	0.144 (0.012)	883	0.151 (0.012)	901	0.145 (0.012)	-0.006	-0.001	0.005
Ward group 4	900	0.122 (0.011)	883	0.123 (0.011)	901	0.115 (0.011)	-0.001	0.007	0.008
Ward group 5	900	0.204 (0.013)	883	0.195 (0.013)	901	0.193 (0.013)	0.010	0.011	0.002
Ward group 6	900	0.218 (0.014)	883	0.203 (0.014)	901	0.204 (0.013)	0.015	0.014	-0.001
Distance to most preferred service (km)	900	1.536 (0.060)	883	1.560 (0.059)	901	1.625 (0.062)	-0.024	-0.089	-0.065
Distance to least preferred service (km)	900	1.705 (0.051)	883	1.654 (0.050)	901	1.682 (0.052)	0.050	0.023	-0.027
Most preferred service: water	900	0.508 (0.017)	883	0.519 (0.017)	901	0.501 (0.017)	-0.011	0.007	0.018
Least preferred service: water	900	0.069 (0.008)	883	0.077 (0.009)	901	0.100 (0.010)	-0.008	-0.031**	-0.023*
Property value (SLL)	900	1.30e+07 (5.80e+05)	883	1.30e+07 (6.06e+05)	901	1.33e+07 (5.94e+05)	-8446.520	-3.51e+05	-3.43e+05

Notes: The value displayed for t-tests are the differences in the means across the groups. ***, **, and * indicate significance at the 1, 5, and 10 percent critical level.

Table A6: Balance of baseline service preferences and distance to services across treatment groups

Variable	N	(1) Match Mean/SE	N	(2) Mismatch Mean/SE	N	(3) Control Mean/SE	(1)-(2)	T-test Difference (1)-(3)	(2)-(3)
most = sanitation	1829	0.093 (0.007)	1831	0.092 (0.007)	1834	0.089 (0.007)	0.002	0.004	0.002
most = health	1829	0.146 (0.008)	1831	0.149 (0.008)	1834	0.148 (0.008)	-0.003	-0.002	0.001
most = markets	1829	0.020 (0.003)	1831	0.031 (0.004)	1834	0.026 (0.004)	-0.010**	-0.005	0.005
most = education	1829	0.036 (0.004)	1831	0.040 (0.005)	1834	0.039 (0.005)	-0.004	-0.003	0.001
most = water	1829	0.537 (0.012)	1831	0.522 (0.012)	1834	0.519 (0.012)	0.015	0.019	0.004
most = environment	1829	0.008 (0.002)	1831	0.010 (0.002)	1834	0.011 (0.002)	-0.003	-0.004	-0.001
most = disaster	1829	0.009 (0.002)	1831	0.008 (0.002)	1834	0.014 (0.003)	0.001	-0.005	-0.007*
most = roads	1829	0.150 (0.008)	1831	0.149 (0.008)	1834	0.154 (0.008)	0.001	-0.004	-0.005
least = sanitation	1829	0.083 (0.006)	1831	0.081 (0.006)	1834	0.081 (0.006)	0.001	0.002	0.001
least = health	1829	0.077 (0.006)	1831	0.072 (0.006)	1834	0.077 (0.006)	0.005	-0.000	-0.005
least = markets	1829	0.083 (0.006)	1831	0.085 (0.007)	1834	0.091 (0.007)	-0.002	-0.008	-0.006
least = education	1829	0.068 (0.006)	1831	0.063 (0.006)	1834	0.059 (0.005)	0.006	0.009	0.004
least = water	1829	0.077 (0.006)	1831	0.082 (0.006)	1834	0.092 (0.007)	-0.005	-0.015	-0.010
least = environment	1829	0.267 (0.010)	1831	0.264 (0.010)	1834	0.260 (0.010)	0.002	0.007	0.004
least = disaster	1829	0.215 (0.010)	1831	0.220 (0.010)	1834	0.213 (0.010)	-0.005	0.002	0.007
least = roads	1829	0.131 (0.008)	1831	0.133 (0.008)	1834	0.128 (0.008)	-0.002	0.003	0.005
Approximate walking distance in km: sanitation	1829	1.439 (0.026)	1831	1.469 (0.027)	1834	1.486 (0.028)	-0.031	-0.047	-0.017
Approximate walking distance in km: health	1829	3.182 (0.037)	1831	3.191 (0.037)	1834	3.190 (0.037)	-0.009	-0.008	0.002
Approximate walking distance in km: markets	1829	2.297 (0.053)	1831	2.331 (0.057)	1834	2.335 (0.057)	-0.035	-0.039	-0.004
Approximate walking distance in km: education	1829	2.991 (0.044)	1831	3.004 (0.045)	1834	3.004 (0.046)	-0.013	-0.014	-0.000
Approximate walking distance in km: water	1829	0.486 (0.008)	1831	0.497 (0.008)	1834	0.484 (0.008)	-0.011	0.003	0.014
Approximate walking distance in km: environment	1829	1.232 (0.020)	1831	1.253 (0.021)	1834	1.250 (0.020)	-0.020	-0.018	0.002
Approximate walking distance in km: disaster	1829	0.908 (0.017)	1831	0.939 (0.019)	1834	0.938 (0.019)	-0.030	-0.030	0.000
Approximate walking distance in km: roads	1829	3.577 (0.073)	1831	3.665 (0.075)	1834	3.641 (0.075)	-0.088	-0.064	0.024
Approximate walking distance in km: most	1829	1.518 (0.042)	1831	1.537 (0.041)	1834	1.515 (0.041)	-0.019	0.003	0.022
Approximate walking distance in km: least	1829	1.644 (0.035)	1831	1.681 (0.036)	1834	1.643 (0.035)	-0.038	0.001	0.039
Above median distance (most)	1829	0.483 (0.012)	1831	0.487 (0.012)	1834	0.481 (0.012)	-0.004	0.002	0.006
Above median distance (least)	1829	0.484 (0.012)	1831	0.507 (0.012)	1834	0.495 (0.012)	-0.023	-0.010	0.013

Notes: The value displayed for t-tests are the differences in the means across the groups. ***, **, and * indicate significance at the 1, 5, and 10 percent critical level.

1.B   Regression Results

Table B1: Treatment effects on the amount of tax paid including different control variables (September 2022)

	Amount paid (SLE)			
	(1)	(2)	(3)	(4)
Panel A. Overall treatment effect				
Treated	23.784**	22.880**	17.946*	22.114**
	(11.322)	(10.993)	(9.799)	(9.627)
Panel B. Treatment effect by arm				
Match	17.226	15.324	4.064	7.894
	(14.639)	(14.378)	(10.222)	(9.857)
Mismatch	31.115**	30.503**	30.568**	35.413**
	(14.913)	(14.704)	(14.390)	(14.277)
Service type FEs (m, l)	No	Yes	Yes	Yes
Distance linear (m, l)	No	Yes	Yes	Yes
Perceived availabilty at BL (m, l)	No	No	Yes	Yes
Wealth FEs	No	No	No	Yes
p-val. Match = Mismatch	0.431	0.393	0.076	0.057
Obs.	5,494	5,494	5,469	5,469
Control mean	118.829	118.829	118.829	118.829
SD	418.244	418.244	418.244	418.244

Notes: Robust standard errors in parentheses. Significance levels are indicated by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Regressions include strata, enumerator, and ward fixed effects and control for respondent age, multiple property use, and the respective outcome variable at baseline.

Table B2: Treatment effects on any payment made including different control variables (September 2022)

	Paid any tax			
	(1)	(2)	(3)	(4)
Panel A. Overall treatment effect				
Treated	0.012	0.012	0.012	0.013
	(0.008)	(0.008)	(0.008)	(0.008)
Panel B. Treatment effect by arm				
Match	0.012	0.011	0.011	0.012
	(0.009)	(0.009)	(0.009)	(0.009)
Mismatch	0.013	0.013	0.014	0.015
	(0.009)	(0.009)	(0.009)	(0.009)
Service type FEs (m, l)	No	Yes	Yes	Yes
Distance linear (m, l)	No	Yes	Yes	Yes
Perceived availabilty at BL (m, l)	No	No	Yes	Yes
Wealth FEs	No	No	No	Yes
p-val. Match = Mismatch	0.949	0.866	0.817	0.778
Obs.	5,494	5,494	5,469	5,469
Control mean	0.201	0.201	0.201	0.201
SD	0.401	0.401	0.401	0.401

Notes: Robust standard errors in parentheses. Significance levels are indicated by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Regressions include strata, enumerator, and ward fixed effects and control for respondent age, multiple property use, and the respective outcome variable at baseline.

Table B3: Treatment effects on the amount of tax paid conditional on payment including different control variables (September 2022)

	Amount paid (SLE, conditional)			
	(1)	(2)	(3)	(4)
Panel A. Overall treatment effect				
Treated	74.983	65.295	54.497	84.546**
	(55.252)	(51.904)	(49.515)	(42.135)
Panel B. Treatment effect by arm				
Match	26.249	2.016	-29.021	13.145
	(64.738)	(60.195)	(51.647)	(43.045)
Mismatch	110.366	104.502	121.975*	143.640**
	(74.488)	(73.076)	(69.479)	(56.784)
Service type FEs (m, l)	No	Yes	Yes	Yes
Distance linear (m, l)	No	Yes	Yes	Yes
Perceived availabilty at BL (m, l)	No	No	Yes	Yes
Wealth FEs	No	No	No	Yes
p-val. Match = Mismatch	0.288	0.190	0.027	0.011
Obs.	1,159	1,159	1,151	1,151
Control mean	590.605	590.605	590.605	590.605
SD	769.365	769.365	769.365	769.365

Notes: Robust standard errors in parentheses. Significance levels are indicated by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Regressions include strata, enumerator, and ward fixed effects and control for respondent age, multiple property use, and the respective outcome variable at baseline.

Table B4: Treatment effects on the amount of tax paid including different controls for distance (September 2022)

	Amount paid (SLE)			
	(1)	(2)	(3)	(4)
Panel A. Overall treatment effect				
Treated	22.114**	22.303**	21.909**	22.437**
	(9.627)	(9.601)	(9.703)	(9.615)
Panel B. Treatment effect by arm				
Match	7.894	8.360	7.654	8.141
	(9.857)	(9.809)	(9.881)	(9.829)
Mismatch	35.413**	35.341**	35.304**	35.472**
	(14.277)	(14.249)	(14.347)	(14.269)
Distance linear (m, l)	Yes	No	No	No
Above median distance (m, l)	No	Yes	No	No
Distance quartiles (m, l)	No	No	Yes	No
Distance linear (all services)	No	No	No	Yes
p-val. Match = Mismatch	0.057	0.060	0.056	0.057
Obs.	5,469	5,469	5,469	5,469
Control mean	118.829	118.829	118.829	118.829
SD	418.244	418.244	418.244	418.244

Notes: Robust standard errors in parentheses. Significance levels are indicated by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Regressions include strata, enumerator, and ward fixed effects and control for respondent age, multiple property use, and the respective outcome variable at baseline. In addition, regressions control for public service preferences (most, least), linear distance to closest services (most, least), perceived availability of services at baseline (most, least), property value, and property ownership.

Table B5: Treatment effects on any payment made including different controls for distance (September 2022)

	Paid any tax			
	(1)	(2)	(3)	(4)
Panel A. Overall treatment effect				
Treated	0.013	0.013	0.013	0.013
	(0.008)	(0.008)	(0.008)	(0.008)
Panel B. Treatment effect by arm				
Match	0.012	0.011	0.012	0.012
	(0.009)	(0.009)	(0.009)	(0.009)
Mismatch	0.015	0.014	0.014	0.014
	(0.009)	(0.009)	(0.009)	(0.009)
Distance linear (m, l)	Yes	No	No	No
Above median distance (m, l)	No	Yes	No	No
Distance quartiles (m, l)	No	No	Yes	No
Distance linear (all services)	No	No	No	Yes
p-val. Match = Mismatch	0.778	0.773	0.778	0.790
Obs.	5,469	5,469	5,469	5,469
Control mean	0.201	0.201	0.201	0.201
SD	0.401	0.401	0.401	0.401

Notes: Robust standard errors in parentheses. Significance levels are indicated by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Regressions include strata, enumerator, and ward fixed effects and control for respondent age, multiple property use, and the respective outcome variable at baseline. In addition, regressions control for public service preferences (most, least), linear distance to closest services (most, least), perceived availability of services at baseline (most, least), property value, and property ownership.

Table B6: Treatment effects on the amount of tax paid conditional on payment including different controls for distance (September 2022)

	Amount paid (SLE, conditional)			
	(1)	(2)	(3)	(4)
Panel A. Overall treatment effect				
Treated	84.546**	79.746*	81.361*	85.316**
	(42.135)	(41.546)	(42.101)	(42.388)
Panel B. Treatment effect by arm				
Match	13.145	8.646	8.809	13.484
	(43.045)	(42.790)	(43.551)	(42.835)
Mismatch	143.640**	139.160**	142.944**	147.082**
	(56.784)	(55.779)	(56.230)	(57.223)
Distance linear (m, l)	Yes	No	No	No
Above median distance (m, l)	No	Yes	No	No
Distance quartiles (m, l)	No	No	Yes	No
Distance linear (all services)	No	No	No	Yes
p-val. Match = Mismatch	0.011	0.011	0.009	0.009
Obs.	1,151	1,151	1,151	1,151
Control mean	590.605	590.605	590.605	590.605
SD	769.365	769.365	769.365	769.365

Notes: Robust standard errors in parentheses. Significance levels are indicated by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Regressions include strata, enumerator, and ward fixed effects and control for respondent age, multiple property use, and the respective outcome variable at baseline. In addition, regressions control for public service preferences (most, least), linear distance to closest services (most, least), perceived availability of services at baseline (most, least), property value, and property ownership.

Figure B1: Distance to treatment service
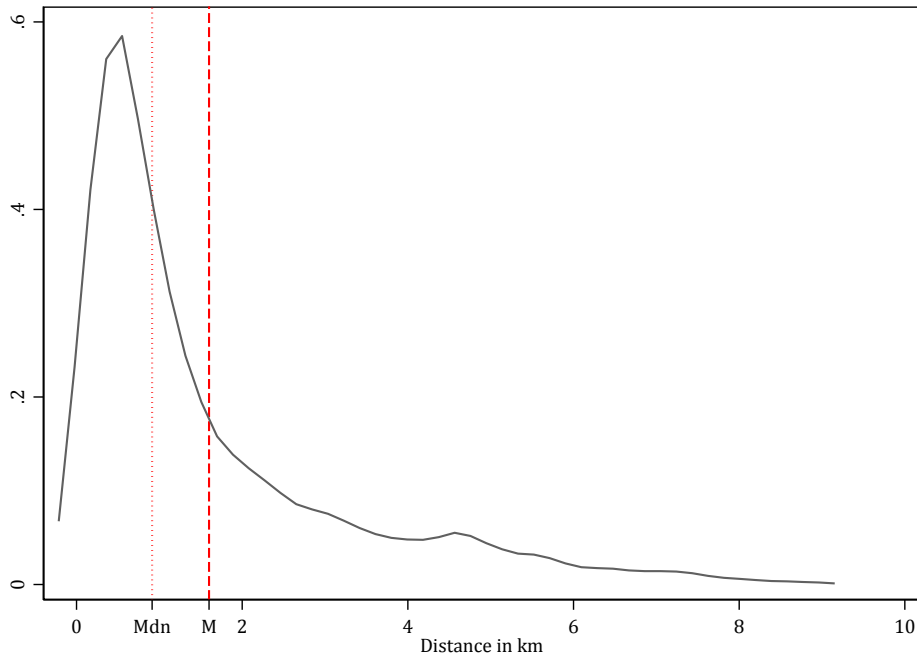
Notes: The figure shows the distribution of distance to the service treated individuals were informed about. Mdn marks the median distance, M marks the mean distance.

Figure B2: Minimum distance to services by type
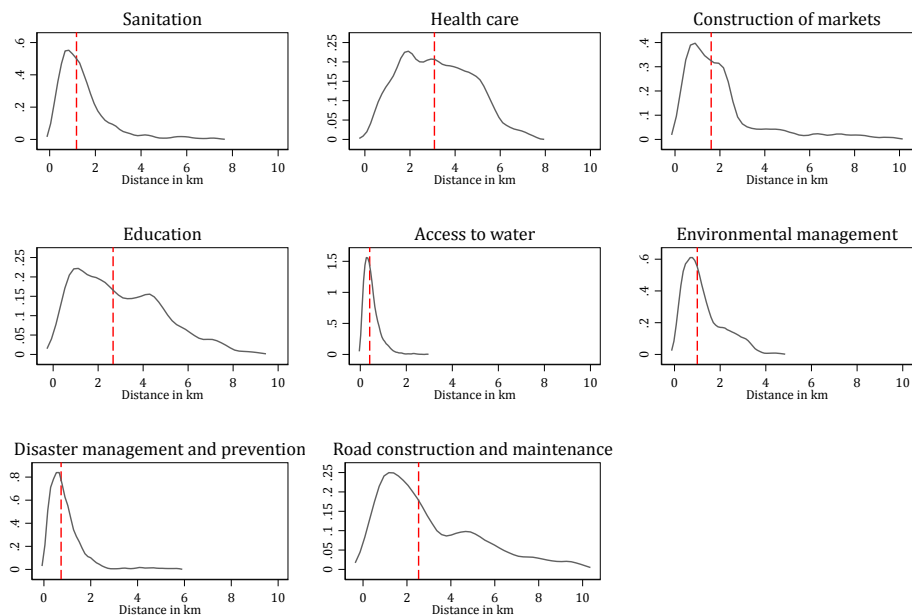
Note: Red lines mark type-specific median distances.

Notes: The figure shows the distribution of distance to the closest service of each type. The dotted line marks the median distance.

Figure B3: Heterogeneity of the treatment effects on whether any tax was paid depending on the content of the treatment message (September 2022)

Figure B4: Heterogeneity of the treatment effects on the amount of tax paid conditional on payment depending on the content of the treatment message (September 2022)
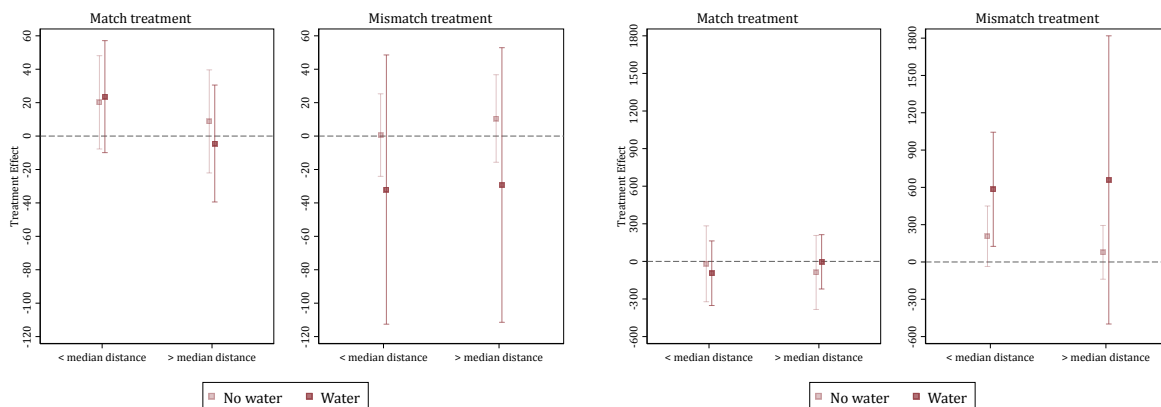
(a) Below median property value   (b) Above median property value

Table B7: Robustness checks—Treatment effects on tax compliance (September 2022)

|	Tax compliance by September 2022						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Amount paid (SLE) (September)	23.565**	23.784**	21.464*	22.114**	20.127**	19.717**	20.459**
	(11.826)	(11.322)	(11.154)	(9.627)	(9.558)	(9.796)	(10.016)
Amount paid (SLE) (win 95, share) (September)	22.824*	23.106**	20.865*	21.359**	19.449**	19.031*	19.742**
	(11.769)	(11.255)	(11.084)	(9.551)	(9.480)	(9.714)	(9.940)
Paid any tax (d) (September)	0.013	0.012	0.011	0.013	0.012	0.012	0.013
	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)
Amount paid (SLE, cond.) (September)	90.164*	74.983	95.681*	84.546**	90.001**	87.242**	86.058*
	(53.291)	(55.252)	(56.998)	(42.135)	(42.597)	(44.095)	(44.338)
Amount paid (SLE, cond.) (win 95, share) (September)	86.623	72.431	93.055	82.479**	87.614**	84.714*	83.501*
	(53.024)	(54.873)	(56.643)	(41.634)	(42.136)	(43.716)	(44.003)
Base controls		Yes	Yes	Yes	Yes	Yes	Yes
Indiv. controls			Yes		Yes	Yes	Yes
Service type FEs (m, l)				Yes	Yes	Yes	Yes
Distance linear (m, l)				Yes	Yes	Yes	Yes
Perceived availabilty at BL (m, l)				Yes	Yes	Yes	Yes
Wealth FEs				Yes	Yes	Yes	Yes
Obs.	1,159	1,159	1,159	1,151	1,151	1,151	1,151
Control mean	587.467	587.467	587.467	587.467	587.467	587.467	587.467
SD	762.420	762.420	762.420	762.420	762.420	762.420	762.420

Notes: Robust standard errors in parentheses. Significance levels are indicated by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Column (1) shows results controlling only for the baseline level of the respective outcome variable. Columns (2) - (7) include strata, enumerator, and ward fixed effects and control for respondent age, multiple property use. Column (2) shows results from the main specification. Column (3) additionally controls for respondent gender, education level, marital status, and income level. Column (4) is based on the main specification and additionally controls for public service preferences (most, least), linear distance to closest services (most, least), perceived availability of services at baseline (most, least), property value, and property ownership. Column (5) combines Columns (3) and (4). Column (6) uses the same controls as Column (4), but does not employ any weights. Column (7) uses the same controls as Column (5) and employs weights adjusting for over-sampling of property owners.

Table B8: Robustness checks—Heterogeneity of the treatment effects on the amount of tax paid conditional on payment depending on the content of the treatment message (September 2022)

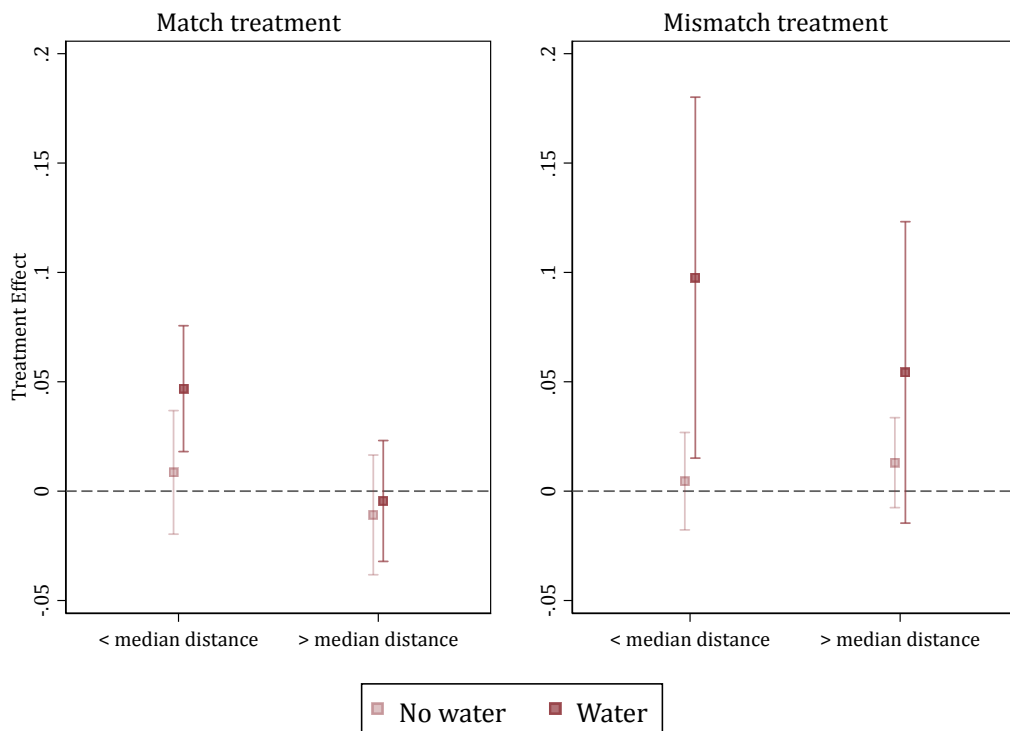|	Any tax paid by September 2022						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Match	0.009	0.011	0.010	0.009	0.008	0.006	0.008
	(0.018)	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)
Match × Water (m)	0.034	0.036	0.037	0.038	0.039	0.040*	0.040*
	(0.024)	(0.024)	(0.024)	(0.024)	(0.024)	(0.024)	(0.024)
Match × > median distance (m)	-0.026	-0.026	-0.025	-0.019	-0.020	-0.019	-0.021
	(0.023)	(0.023)	(0.023)	(0.023)	(0.023)	(0.023)	(0.023)
Match × Water (m) × > median distance (m)	-0.016	-0.025	-0.026	-0.032	-0.031	-0.032	-0.031
	(0.033)	(0.033)	(0.033)	(0.033)	(0.033)	(0.033)	(0.033)
Mismatch	0.008	0.004	0.005	0.005	0.006	0.006	0.008
	(0.014)	(0.014)	(0.014)	(0.014)	(0.013)	(0.013)	(0.013)
Mismatch × Water (l)	0.087*	0.091*	0.087*	0.093*	0.089*	0.089*	0.086*
	(0.051)	(0.052)	(0.051)	(0.052)	(0.051)	(0.051)	(0.051)
Mismatch × > median distance (l)	0.004	0.007	0.003	0.008	0.005	0.004	0.001
	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)
Mismatch × Water (l) × > median distance (l)	-0.046	-0.045	-0.037	-0.052	-0.045	-0.044	-0.043
	(0.066)	(0.068)	(0.067)	(0.068)	(0.067)	(0.067)	(0.067)
Base controls		Yes	Yes	Yes	Yes	Yes	Yes
Indiv. controls		Yes			Yes	Yes	Yes
Service type FEs (m, l)			Yes	Yes	Yes	Yes	Yes
Distance linear (m, l)				Yes	Yes	Yes	Yes
Perceived availabilty at BL (m, l)				Yes	Yes	Yes	Yes
Wealth FEs				Yes	Yes	Yes	Yes
p-val. Match = Mismatch	0.945	0.694	0.780	0.838	0.927	0.975	0.980
Match + water (m) = 0	0.014	0.007	0.007	0.007	0.008	0.008	0.007
Mismatch + water (l) = 0	0.054	0.058	0.064	0.052	0.056	0.056	0.058
Match + > median distance (m) = 0	0.319	0.394	0.364	0.514	0.463	0.428	0.425
Mismatch + > median distance (l) = 0	0.335	0.411	0.517	0.297	0.388	0.428	0.453
Match + water (m) + > median distance (m) = 0	0.932	0.847	0.811	0.789	0.782	0.776	0.761
Mismatch + water (l) + > median distance (l) = 0	0.193	0.170	0.167	0.195	0.195	0.187	0.201
Obs.	5,494	5,494	5,494	5,469	5,469	5,469	5,469
Control mean	0.201	0.201	0.201	0.201	0.201	0.201	0.201
SD	0.401	0.401	0.401	0.401	0.401	0.401	0.401

Notes: Robust standard errors in parentheses. Significance levels are indicated by $* p < 0.10$, $** p < 0.05$, $*** p < 0.01$. Column (1) shows results controlling only for the baseline level of the respective outcome variable. Columns (2) - (7) include strata, enumerator, and ward fixed effects and control for respondent age, multiple property use. Column (2) shows results from the main specification. Column (3) additionally controls for respondent gender, education level, marital status, and income level. Column (4) is based on the main specification and additionally controls for public service preferences (most, least), linear distance to closest services (most, least), perceived availability of services at baseline (most, least), property value, and property ownership. Column (5) combines Columns (3) and (4). Column (6) uses the same controls as Column (4), but does not employ any weights. Column (7) uses the same controls as Column (5) and employs weights adjusting for over-sampling of property owners.

Endline Data Measurement Challenges   As explained in Section 1.3, the endline survey was administered to the subsample of respondents containing both property owners and tenants in the proportion in which they were reached during the baseline interview. As opposed to this, in the overall sample, property owners are oversampled. Robustness checks suggest that the tax compliance results are robust to the use of weights adjusting for the share of property owners in the randomly selected subsample of owners and tenants (see Tables B7 and B8). This is important, as scaling up this intervention to a larger set of properties will likely yield a similar distribution between property owners and tenants being contacted, given that phone numbers do not necessarily belong to the owner (alone), but often to a tenant, family member, or caretaker. Based on these results, the self-reported endline information can be used to draw conclusions about potential mechanisms in this population.

An additional challenge, however, is selective attrition within the sample of individuals to be interviewed at endline. As compared to those who participated in the endline survey, attrited individuals are more likely to be female and more educated, have higher monthly income, live in different areas of the city, are more likely to indicate water as their least preferred service, and live in more valuable properties (see Table B9). Whereas the sample is still well-balance across treatment groups, it is not representative for the population of individuals interviewed at baseline (see Table A5).

Despite these differences in characteristics, the tax compliance results presented above largely hold true for the subsample of individuals who participated in the endline interview.[70] This does not guarantee that the channels leading to these effects are the same as for the remainder of the sample, but can serve as indicative evidence for the comparability of both samples with respect to their tax compliance behavior despite differences in observables.

---

[70] See Table B10 for a summary of results using the complete sample and Table B11 for the sample of individuals interviewed at endline only.

Table B9: Baseline characteristics of attrited (left) vs. non-attrited (right) individuals

Variable	N	(1) No Mean/SE	N	(2) Yes Mean/SE	T-test Difference (1)-(2)
Female	705	0.417 (0.019)	2684	0.349 (0.009)	0.068***
Age	701	5.399 (0.096)	2673	5.447 (0.050)	-0.048
Education level	696	4.409 (0.086)	2682	4.208 (0.045)	0.202**
In relationship	702	0.886 (0.012)	2683	0.873 (0.006)	0.013
Owner	705	0.652 (0.018)	2684	0.644 (0.009)	0.008
Monthly income	589	3.635 (0.078)	2577	3.423 (0.036)	0.212**
Ward group 1	705	0.128 (0.013)	2684	0.159 (0.007)	-0.031**
Ward group 2	705	0.106 (0.012)	2684	0.168 (0.007)	-0.062***
Ward group 3	705	0.123 (0.012)	2684	0.147 (0.007)	-0.023
Ward group 4	705	0.119 (0.012)	2684	0.120 (0.006)	-0.001
Ward group 5	705	0.251 (0.016)	2684	0.197 (0.008)	0.054***
Ward group 6	705	0.272 (0.017)	2684	0.208 (0.008)	0.064***
Distance to most preferred service (km)	705	1.587 (0.066)	2684	1.573 (0.035)	0.013
Distance to least preferred service (km)	705	1.658 (0.058)	2684	1.680 (0.030)	-0.023
Most preferred service: water	705	0.504 (0.019)	2684	0.509 (0.010)	-0.005
Least preferred service: water	705	0.106 (0.012)	2684	0.082 (0.005)	0.024**
Property value (SLL)	705	1.73e+07 (8.71e+05)	2684	1.31e+07 (3.42e+05)	4.20e+06***

Notes: The value displayed for t-tests are the differences in the means across the groups. ***, **, and * indicate significance at the 1, 5, and 10 percent critical level.

Table B10: Treatment effects on whether any tax was paid—entire sample

	Any tax paid by September 2022				
	(1)	(2)	(3)	(4)	(5)
Treated	0.012	0.013			
	(0.008)	(0.008)			
Match			0.012	0.012	0.009
			(0.009)	(0.009)	(0.017)
Mismatch			0.013	0.015	0.005
			(0.009)	(0.009)	(0.014)
Match × Water (m)					0.038
					(0.024)
Match × > median distance (m)					-0.019
					(0.023)
Match × Water (m) × > median distance (m)					-0.032
					(0.033)
Mismatch × Water (l)					0.093*
					(0.052)
Mismatch × > median distance (l)					0.008
					(0.017)
Mismatch × Water (l) × > median distance (l)					-0.052
					(0.068)
Service type FEs (m, l)	No	Yes	No	Yes	Yes
Distance linear (m, l)	No	Yes	No	Yes	Yes
Perceived availabilty at BL (m, l)	No	Yes	No	Yes	Yes
Wealth FEs	No	Yes	No	Yes	Yes
p-val. Match = Mismatch			0.949	0.778	0.838
Match + water (m) = 0					0.007
Mismatch + water (l) = 0					0.052
Match + > median distance (m) = 0					0.514
Mismatch + > median distance (l) = 0					0.297
Match + water (m) + > median distance (m) = 0					0.789
Mismatch + water (l) + > median distance (l) = 0					0.195
Obs.	5,494	5,469	5,494	5,469	5,469
Control mean	0.201	0.201	0.201	0.201	0.201
SD	0.401	0.401	0.401	0.401	0.401

Notes: Robust standard errors in parentheses. Significance levels are indicated by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Regressions include strata, enumerator, and ward fixed effects and control for respondent age, multiple property use, and the respective outcome variable at baseline.

Table B11: Treatment effects on whether any tax was paid—endline sample

	Any tax paid by September 2022				
	(1)	(2)	(3)	(4)	(5)
Treated	0.011	0.012			
	(0.011)	(0.011)			
Match			0.020	0.021	0.016
			(0.013)	(0.013)	(0.023)
Mismatch			0.004	0.005	0.008
			(0.013)	(0.013)	(0.019)
Match × Water (m)					0.064*
					(0.034)
Match × > median distance (m)					-0.026
					(0.029)
Match × Water (m) × > median distance (m)					-0.061
					(0.045)
Mismatch × Water (l)					0.086
					(0.070)
Mismatch × > median distance (l)					-0.013
					(0.024)
Mismatch × Water (l) × > median distance (l)					-0.069
					(0.086)
Service type FEs (m, l)	No	Yes	No	Yes	Yes
Distance linear (m, l)	No	Yes	No	Yes	Yes
Perceived availabilty at BL (m, l)	No	Yes	No	Yes	Yes
Wealth FEs	No	Yes	No	Yes	Yes
p-val. Match = Mismatch			0.246	0.230	0.766
Match + water (m) = 0					0.002
Mismatch + water (l) = 0					0.170
Match + > median distance (m) = 0					0.624
Mismatch + > median distance (l) = 0					0.785
Match + water (m) + > median distance (m) = 0					0.760
Mismatch + water (l) + > median distance (l) = 0					0.771
Obs.	2,684	2,664	2,684	2,664	2,664
Control mean	0.201	0.201	0.201	0.201	0.201
SD	0.401	0.401	0.401	0.401	0.401

Notes: Robust standard errors in parentheses. Significance levels are indicated by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Regressions include strata, enumerator, and ward fixed effects and control for respondent age, multiple property use, and the respective outcome variable at baseline.

Table B12: Treatment effects on perceived government performance

	Original	Inverse prob. weights	Horowitz-Manski bounds	
			Upper	Lower
	(1)	(2)	(3)	(4)
Performance of local government index	0.051	0.051	1.129***	-1.021***
	(0.036)	(0.038)	(0.050)	(0.049)
	[2,684]	[2,684]	[3,389]	[3,389]
- Confidence in ward councilor	0.032	0.024	0.727***	-0.648***
	(0.036)	(0.039)	(0.038)	(0.040)
	[2,684]	[2,684]	[3,389]	[3,389]
- Confidence in mayor	0.025	0.019	0.857***	-0.813***
	(0.037)	(0.040)	(0.046)	(0.043)
	[2,684]	[2,684]	[3,389]	[3,389]
- Confidence in FCC	0.050	0.066*	1.005***	-0.886***
	(0.037)	(0.039)	(0.047)	(0.045)
	[2,684]	[2,684]	[3,389]	[3,389]
- Approval of ward councilor	0.050	0.045	0.770***	-0.685***
	(0.036)	(0.039)	(0.039)	(0.041)
	[2,684]	[2,684]	[3,389]	[3,389]
- Approval of mayor	0.028	0.029	0.936***	-0.872***
	(0.035)	(0.037)	(0.049)	(0.045)
	[2,684]	[2,684]	[3,389]	[3,389]
Local government responsiveness (std.)	0.043	0.036	0.803***	-0.683***
	(0.033)	(0.035)	(0.042)	(0.041)
	[2,610]	[2,610]	[3,304]	[3,304]

Notes: Robust standard errors in parentheses. Number of observations in brackets. Significance levels are indicated by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Regressions include strata, enumerator, and ward fixed effects and control for respondent age, multiple property use, and the respective outcome variable at baseline.

Table B13: Treatment effects on perceived government performance (below median value properties)

| | Original | Inverse prob. | Horowitz-Manski bounds | |
| | | weights | Upper | Lower |
	(1)	(2)	(3)	(4)
Performance of local government index	0.057	0.078	0.936***	-0.838***
	(0.054)	(0.056)	(0.071)	(0.070)
	[1,395]	[1,395]	[1,693]	[1,693]
- Confidence in ward councilor	0.041	0.065	0.644***	-0.489***
	(0.055)	(0.058)	(0.055)	(0.059)
	[1,395]	[1,395]	[1,693]	[1,693]
- Confidence in mayor	0.040	0.045	0.696***	-0.681***
	(0.058)	(0.060)	(0.067)	(0.064)
	[1,395]	[1,395]	[1,693]	[1,693]
- Confidence in FCC	0.043	0.057	0.818***	-0.744***
	(0.056)	(0.060)	(0.068)	(0.066)
	[1,395]	[1,395]	[1,693]	[1,693]
- Approval of ward councilor	0.041	0.046	0.646***	-0.555***
	(0.054)	(0.058)	(0.057)	(0.059)
	[1,395]	[1,395]	[1,693]	[1,693]
- Approval of mayor	0.043	0.064	0.751***	-0.741***
	(0.053)	(0.052)	(0.070)	(0.065)
	[1,395]	[1,395]	[1,693]	[1,693]
Local government responsiveness (std.)	0.057	0.062	0.673***	-0.537***
	(0.048)	(0.050)	(0.060)	(0.060)
	[1,368]	[1,368]	[1,659]	[1,659]

Notes: Robust standard errors in parentheses. Number of observations in brackets. Significance levels are indicated by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Regressions include strata, enumerator, and ward fixed effects and control for respondent age, multiple property use, and the respective outcome variable at baseline.

Table B14: Treatment effects on perceived government performance (above median value properties)

| | Original | Inverse prob. | Horowitz-Manski bounds | |
| | | weights | Upper | Lower |
	(1)	(2)	(3)	(4)
Performance of local government index	0.001	-0.011	1.332***	-1.251***
	(0.056)	(0.060)	(0.077)	(0.073)
	[1,289]	[1,289]	[1,696]	[1,696]
- Confidence in ward councilor	-0.000	-0.035	0.815***	-0.837***
	(0.057)	(0.060)	(0.057)	(0.059)
	[1,289]	[1,289]	[1,696]	[1,696]
- Confidence in mayor	-0.017	-0.022	1.027***	-0.979***
	(0.056)	(0.061)	(0.069)	(0.064)
	[1,289]	[1,289]	[1,696]	[1,696]
- Confidence in FCC	0.044	0.059	1.211***	-1.062***
	(0.058)	(0.061)	(0.070)	(0.067)
	[1,289]	[1,289]	[1,696]	[1,696]
- Approval of ward councilor	0.022	0.000	0.891***	-0.857***
	(0.056)	(0.059)	(0.059)	(0.060)
	[1,289]	[1,289]	[1,696]	[1,696]
- Approval of mayor	-0.031	-0.035	1.134***	-1.038***
	(0.053)	(0.058)	(0.074)	(0.068)
	[1,289]	[1,289]	[1,696]	[1,696]
Local government responsiveness (std.)	0.046	0.040	0.960***	-0.857***
	(0.050)	(0.055)	(0.063)	(0.061)
	[1,242]	[1,242]	[1,645]	[1,645]

Notes: Robust standard errors in parentheses. Number of observations in brackets. Significance levels are indicated by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Regressions include strata, enumerator, and ward fixed effects and control for respondent age, multiple property use, and the respective outcome variable at baseline.

Table B15: Treatment effects on perceived availability of and satisfaction with public services

| | Original | Inverse prob. | Horowitz-Manski bounds | |
| | | weights | Upper | Lower |
	(1)	(2)	(3)	(4)
Availability of public services index	0.034	0.034	1.261***	-1.216***
	(0.029)	(0.031)	(0.052)	(0.055)
	[2,684]	[2,684]	[3,389]	[3,389]
Availability of most preferred service (std.)	0.037	0.054	0.826***	-0.741***
	(0.038)	(0.040)	(0.041)	(0.044)
	[2,677]	[2,677]	[3,385]	[3,385]
Availability of least preferred service (std.)	-0.005	-0.005	0.772***	-0.770***
	(0.038)	(0.040)	(0.041)	(0.043)
	[2,665]	[2,665]	[3,371]	[3,371]
Satisfaction with public services index	0.058**	0.057*	1.106***	-0.955***
	(0.029)	(0.031)	(0.047)	(0.048)
	[2,683]	[2,683]	[3,388]	[3,388]
Satisfaction with most preferred service (std.)	0.065*	0.090**	0.753***	-0.627***
	(0.037)	(0.040)	(0.039)	(0.040)
	[2,676]	[2,676]	[3,384]	[3,384]
Satisfaction with least preferred service (std.)	0.071*	0.064*	0.738***	-0.589***
	(0.037)	(0.039)	(0.039)	(0.039)
	[2,669]	[2,669]	[3,376]	[3,376]

Notes: Robust standard errors in parentheses. Number of observations in brackets. Significance levels are indicated by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Regressions include strata, enumerator, and ward fixed effects and control for respondent age, multiple property use, and the respective outcome variable at baseline.

Table B16: Treatment effects on satisfaction with local level public service provision (below median value properties)

	(1)	(2)	(3)
	Satisfaction with public services index	Satisfaction with most preferred service (std.)	Satisfaction with least preferred service (std.)
Panel A. Overall treatment effect			
Treated	0.079*	0.074	0.043
	(0.044)	(0.054)	(0.052)
Panel B. Treatment effect by arm			
Match	0.071	0.090	0.062
	(0.050)	(0.064)	(0.060)
Mismatch	0.088*	0.052	0.019
	(0.053)	(0.063)	(0.063)
p-val. Match = Mismatch	0.737	0.526	0.468
Obs.	1,385	1,382	1,382
Control mean	-0.000	-0.000	0.000
SD	1.000	1.000	1.000

Notes: Robust standard errors in parentheses. Significance levels are indicated by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Regressions include strata, enumerator, and ward fixed effects and control for respondent age, multiple property use, and the respective outcome variable at baseline. In addition, regressions control for public service preferences (most, least), linear distance to closest services (most, least), perceived availability of services at baseline (most, least), property value, and property ownership.

Figure B5: Heterogeneity of the treatment effects on whether any payment was made depending on the content of the treatment message (December 2022)
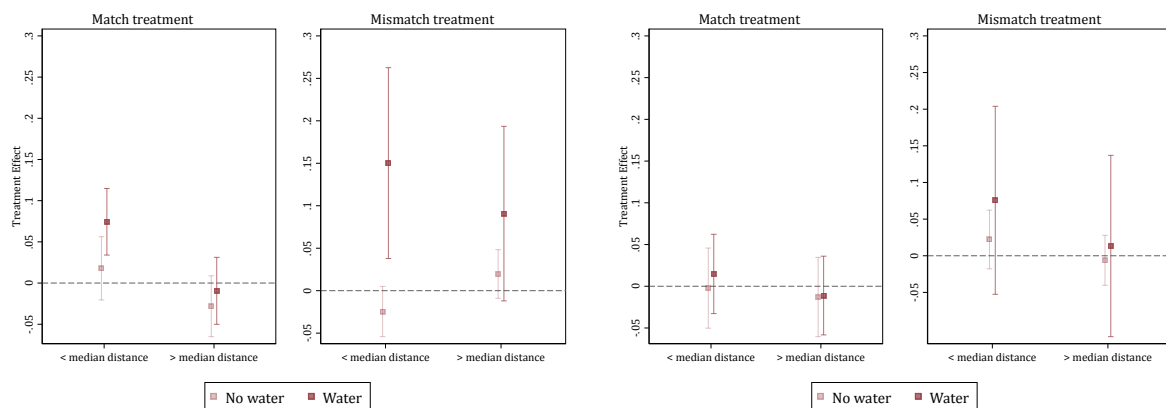
(a) Below median property value   (b) Above median property value

Table B17: Treatment effects on satisfaction with local level public service provision (above median value properties)

	(1) Satisfaction with public services index	(2) Satisfaction with most preferred service (std.)	(3) Satisfaction with least preferred service (std.)
Panel A. Overall treatment effect			
Treated	0.044	0.038	0.101*
	(0.041)	(0.053)	(0.054)
Panel B. Treatment effect by arm			
Match	0.030	0.036	0.058
	(0.049)	(0.064)	(0.064)
Mismatch	0.063	0.042	0.163**
	(0.049)	(0.062)	(0.064)
p-val. Match = Mismatch	0.499	0.927	0.092
Obs.	1,279	1,278	1,277
Control mean	-0.000	-0.000	0.000
SD	1.000	1.000	1.000

Notes: Robust standard errors in parentheses. Significance levels are indicated by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Regressions include strata, enumerator, and ward fixed effects and control for respondent age, multiple property use, and the respective outcome variable at baseline. In addition, regressions control for public service preferences (most, least), linear distance to closest services (most, least), perceived availability of services at baseline (most, least), property value, and property ownership.

Table B18: Heterogeneous treatment effects by perceived availability at baseline

Dep. var.: Any tax paid by September 2022	(1)
Match	0.021
	(0.013)
A little/lot (m)	0.009
	(0.012)
Match × A little/lot (m)	-0.017
	(0.017)
Mismatch	0.010
	(0.014)
A little/lot (l)	0.013
	(0.012)
Mismatch × A little/lot (l)	0.007
	(0.018)
p-val. Match = Mismatch	0.474
p-val. Match + Match × Perceived availability (m) = 0	0.730
p-val. Mismatch + Mismatch × Perceived availability (l) = 0	0.159
Obs.	5,469
Control mean	0.201
SD	0.401

Notes: Robust standard errors in parentheses. Significance levels are indicated by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Regressions include strata, enumerator, and ward fixed effects and control for respondent age, multiple property use, and the respective outcome variable at baseline. In addition, regressions control for public service preferences (most, least), linear distance to closest services (most, least), property value, and property ownership.

Table B19: Treatment effects on perceived government information on taxpayers

| | Original | Inverse prob. | Horowitz-Manski bounds | |
| | | weights | Upper | Lower |
	(1)	(2)	(3)	(4)
Information on taxpayers index	0.035	0.043	1.755***	-1.644***
	(0.037)	(0.040)	(0.070)	(0.066)
	[2,683]	[2,683]	[3,387]	[3,387]
HH location	-0.013	0.001	1.389***	-1.351***
	(0.032)	(0.037)	(0.079)	(0.063)
	[2,683]	[2,683]	[3,387]	[3,387]
Compliance	0.048	0.052*	0.701***	-0.566***
	(0.029)	(0.030)	(0.041)	(0.039)
	[2,683]	[2,683]	[3,387]	[3,387]
Occupation	0.077**	0.074**	0.744***	-0.589***
	(0.035)	(0.037)	(0.040)	(0.045)
	[2,683]	[2,683]	[3,387]	[3,387]
Income	-0.038	-0.039	0.738***	-0.841***
	(0.031)	(0.033)	(0.043)	(0.049)
	[2,683]	[2,683]	[3,387]	[3,387]

Notes: Robust standard errors in parentheses. Number of observations in brackets. Significance levels are indicated by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Regressions include strata, enumerator, and ward fixed effects and control for respondent age, multiple property use, and the respective outcome variable at baseline.

Table B20: Heterogeneous treatment effects on tax compliance by informal public good contribution (September 2022)

	Amount paid (SLE)		Paid any tax (d)		Amount paid (SLE, cond.)	
	(1) Did not pay	(2) Paid	(3) Did not pay	(4) Paid	(5) Did not pay	(6) Paid
Treated	30.815	20.326**	-0.044*	0.019**	56.633	72.092
	(41.345)	(9.793)	(0.026)	(0.009)	(242.370)	(45.995)
p-val T S1=S2	0.093		0.051		0.247	
Match	-19.921	13.739	-0.048	0.019*	-377.200	19.028
	(40.313)	(10.742)	(0.031)	(0.010)	(468.303)	(46.805)
Mismatch	74.746	24.878*	-0.037	0.019*	241.131	115.693*
	(55.623)	(13.840)	(0.030)	(0.010)	(281.165)	(62.255)
p-val. Match = Mismatch	0.047	0.438	0.710	0.967	0.134	0.087
p-val. Match S1 = S2	0.177		0.075		0.502	
p-val. Mismatch S1 = S2	0.172		0.126		0.200	
Obs.	835	4,623	835	4,623	201	947
Control mean	150.260	111.937	0.249	0.191	603.125	584.729
SD	509.186	398.036	0.433	0.394	880.174	743.296

Notes: Robust standard errors in parentheses. Significance levels are indicated by $* p < 0.10$, $** p < 0.05$, $*** p < 0.01$. Regressions include strata, enumerator, and ward fixed effects and control for respondent age, multiple property use, and the respective outcome variable at baseline. In addition, regressions control for public service preferences (most, least), linear distance to closest services (most, least), perceived availability of services at baseline (most, least), property value, and property ownership.

Table B21: Heterogeneity of the treatment effects on perceived extractive capacity of the government depending on the content of the treatment message

	(1) Information on taxpayers index	(2) HH location	(3) Compliance	(4) Occupation	(5) Income
Match	0.073	0.052	0.040	0.047	0.012
	(0.069)	(0.052)	(0.059)	(0.071)	(0.056)
Match × > median distance (m)	-0.033	-0.052	0.036	0.055	-0.106
	(0.102)	(0.092)	(0.079)	(0.104)	(0.082)
Match × Water (m)	-0.040	-0.030	0.035	-0.087	-0.005
	(0.099)	(0.081)	(0.075)	(0.096)	(0.081)
Match × > median distance (m) × Water (m)	0.025	0.028	-0.203*	0.162	0.069
	(0.149)	(0.133)	(0.113)	(0.148)	(0.118)
Mismatch	-0.009	-0.024	0.041	0.039	-0.075
	(0.054)	(0.046)	(0.044)	(0.055)	(0.048)
Mismatch × > median distance (l)	0.074	0.008	0.027	0.086	0.032
	(0.074)	(0.070)	(0.060)	(0.076)	(0.062)
Mismatch × Water (l)	0.126	0.339**	-0.089	-0.043	0.074
	(0.182)	(0.160)	(0.151)	(0.174)	(0.107)
Mismatch × > median distance (l) × Water (l)	0.200	-0.242	0.251	0.114	0.258
	(0.250)	(0.194)	(0.201)	(0.258)	(0.207)
p-val. Match = Mismatch	0.240	0.184	0.984	0.910	0.144
p-val. Match + Match × > median distance (m) = 0	0.615	0.998	0.190	0.210	0.148
p-val. Match + Mismatch + Mismatch × > median distance (l) = 0	0.293	0.795	0.152	0.037	0.390
p-val. Match + Match × Water (m) = 0	0.672	0.734	0.159	0.573	0.917
p-val. Match + Mismatch + Mismatch × Water (l) = 0	0.506	0.039	0.743	0.981	0.998
p-val. Match + > median distance (m) + Water (m) = 0	0.751	0.980	0.165	0.031	0.619
p-val. Mismatch + > median distance (l) + Water (l) = 0	0.014	0.376	0.065	0.274	0.089
Obs.	2,663	2,663	2,663	2,663	2,663
Control mean	0.000	-0.000	0.000	0.000	0.000
SD	1.000	1.000	1.000	1.000	1.000

Notes: Robust standard errors in parentheses. Significance levels are indicated by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Regressions include strata, enumerator, and ward fixed effects and control for respondent age, multiple property use, and the respective outcome variable at baseline. In addition, regressions control for public service preferences (most, least), linear distance to closest services (most, least), perceived availability of services at baseline (most, least), property value, and property ownership.

Table B22: Treatment effects on perceived extractive capacity of the government (above median value properties)

| | Original | Inverse prob. weights | Horowitz-Manski bounds | |
| | | | Upper | Lower |
	(1)	(2)	(3)	(4)
Information on taxpayers index	0.040	0.067	2.112***	-1.978***
	(0.056)	(0.060)	(0.108)	(0.103)
	[1,288]	[1,288]	[1,695]	[1,695]
HH location	-0.035	-0.016	1.716***	-1.579***
	(0.051)	(0.056)	(0.121)	(0.099)
	[1,288]	[1,288]	[1,695]	[1,695]
Compliance	0.053	0.076	0.837***	-0.690***
	(0.047)	(0.049)	(0.061)	(0.059)
	[1,288]	[1,288]	[1,695]	[1,695]
Occupation	0.058	0.058	0.839***	-0.765***
	(0.054)	(0.058)	(0.060)	(0.067)
	[1,288]	[1,288]	[1,695]	[1,695]
Income	0.008	0.021	0.908***	-0.991***
	(0.046)	(0.049)	(0.066)	(0.073)
	[1,288]	[1,288]	[1,695]	[1,695]

Notes: Robust standard errors in parentheses. Number of observations in brackets. Significance levels are indicated by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Regressions include strata, enumerator, and ward fixed effects and control for respondent age, multiple property use, and the respective outcome variable at baseline.

Table B23: Treatment effects on perceived extractive capacity of the government by treatment arm (above median value properties)

	(1) Information on taxpayers index	(2) HH location	(3) Compliance	(4) Occupation	(5) Income
Panel A. Original					
Match	0.051	-0.023	0.049	0.060	0.021
	(0.067)	(0.058)	(0.055)	(0.065)	(0.053)
Mismatch	0.026	-0.050	0.060	0.066	-0.022
	(0.066)	(0.064)	(0.055)	(0.067)	(0.057)
p-val. Match = Mismatch	0.745	0.734	0.511	0.516	0.769
Obs.	1,278	1,278	1,278	1,278	1,278
Panel B. Inverse probability weights					
Match	0.052	-0.018	0.061	0.032	0.033
	(0.072)	(0.065)	(0.057)	(0.071)	(0.056)
Mismatch	0.079	-0.011	0.093	0.092	-0.010
	(0.068)	(0.069)	(0.057)	(0.069)	(0.059)
p-val. Match = Mismatch	0.500	0.963	0.259	0.408	0.740
Obs.	1,278	1,278	1,278	1,278	1,278

Notes: Robust standard errors in parentheses. Significance levels are indicated by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Regressions include strata, enumerator, and ward fixed effects and control for respondent age, multiple property use, and the respective outcome variable at baseline. In addition, regressions control for public service preferences (most, least), linear distance to closest services (most, least), perceived availability of services at baseline (most, least), property value, and property ownership.

Table B24: Treatment effects on tax compliance (January – June 2023)

	(1) Amount paid (SLE)	(2) Paid any tax (d)	(3) Amount paid (SLE, cond.)
Treated	-5.668	0.003	-55.269
	(17.451)	(0.009)	(101.327)
Obs.	5,494	5,494	833
Control mean	134.538	0.149	900.524
SD	613.328	0.357	1354.032

Notes: Robust standard errors in parentheses. Significance levels are indicated by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Regressions include strata, enumerator, and ward fixed effects and control for respondent age, multiple property use, and the respective outcome variable at baseline.

Chapter 2

The Power of Faith: Effects of an Imam-led Information Campaign on Labor Supply and Social Interactions

Joint work with Alexandra Avdeenko, Jakob Gärtner, Marc Gillaizeau, Ghida Karbala, Giulia Montresor, Atika Pasha, and Galina Zudenkova

Abstract: We conduct a field experiment randomizing a remote awareness campaign that promotes social distancing and preventive health behaviors to contain disease spread in rural Pakistan. We examine the impact caused by religious leaders' endorsements of the campaign on labor supply and social interaction patterns. Our results show that the campaign led to reductions in labor supply and social interactions only when accompanied by Imams' public endorsements. The effects are driven by male individuals, who adjust behavior in response to greater concerns about the risk of transmitting a disease to others. Despite increases in knowledge, we observe no effects on behavior among female individuals. Our results highlight the informational role of religious leaders in shaping critical beliefs and behavior in the context of a health crisis.

2.1 Introduction

The effectiveness of information campaigns in shaping individual behavior hinges on the credibility and persuasiveness of the source of information. This is especially important in times of crises such as pandemics, when information is spread through a variety of channels, and individuals have to decide whom to trust. In many developing countries, skepticism towards governmental and health authorities undermines their ability to guide individuals' decisions on the adoption of health-beneficial behavior (Christensen et al., 2021). Religious institutions, on the other hand, enjoy widespread trust and have substantial influence over their adherents' beliefs and behavior (Stroebel and van Benthem, 2012; Bassi and Rasul, 2017; Condra et al., 2019; Moreno-Medina, 2023). Their role may be of particular importance during times of hardship, when the number of individuals practicing their religion increases as people seek solace in faith (Belloc et al., 2016; Bentzen, 2021).

In this study, we investigate whether information spread by religious leaders can effectively change individuals' attitudes towards health risks and their compliance with recommended health measures. We conduct a randomized controlled trial with NGO-beneficiaries from 886 rural villages in Pakistan, testing the effectiveness of a remote awareness campaign and its endorsement by religious leaders in the context of the COVID-19 pandemic. We randomly allocate each village to one of three experimental groups. Individuals in the respective villages either receive no information at all (control group), an information call by representatives of the NGO (*phone* group), or an information call by representatives of the NGO as well as loudspeaker announcements by Imams in their village, endorsing the content of the information call (*phone plus loudspeaker* group). During the awareness call, individuals receive information about the characteristics of the virus, as well as potential measures to protect oneself and others against an infection. The loudspeaker announcements are an abbreviated version of the call, summarizing the most important information.

We test the effects of the two treatments as compared to the control group, as well as compared to each other, on three main outcomes of interest: labor supply, social interactions, and the use of preventive measures. To get a better understanding of the channels leading to changes along these dimensions, we further look at the effects of the treatments on knowledge about the virus, as well as beliefs related to the probabilities of getting infected or infecting others, the severity of the virus, and the cost related to an infection. We measure these outcomes making use of information gathered during our endline survey, approximately two months after the treatment.

Our results show that the phone campaign endorsed by religious leaders via mosque loudspeakers led to a significant drop in labor supply. Individuals assigned to the *phone plus loudspeaker* treatment are 4.2 percentage points less likely to have worked outside home in the seven

days prior to the endline interview. The effect is driven by male individuals, who are ex ante substantially more likely to work than women, in particular outside home (50% versus 26%, respectively). In response to the treatment, their likelihood of having worked outside home decreases by 4.8 percentage points, and they work about half a workday per week less on average. We observe no changes in labor activity among women. Similarly, we find that male individuals in the *phone plus loudspeaker* group significantly decrease social interactions. We detect no changes in the use of preventive measures such as mask-wearing or hand-washing.

The effects of the awareness campaign endorsed by religious leaders on labor supply and social interactions differ significantly from those of the awareness campaign alone, for which we detect no significant effects on behavior. This suggests that the campaign endorsement by religious leaders was indeed critical in generating the observed behavioral adjustments. In response to concerns that additional messages distributed by Imams may affect outcomes through a reminder effect rather than through the credibility of the messenger, we provide additional suggestive evidence in support of our hypothesis. First, we find no evidence that differences in treatment intensity within the *phone plus loudspeaker* group affect outcomes. Second, we show that the effects of the treatment differ along the intensive margin of religiosity. Finally, information from our endline survey indicates a significant presence of information campaigns (implemented via both phone and loudspeakers) in our study area, reaching individuals across all experimental groups. However, individuals in the *phone plus loudspeaker* group are substantially more likely to report that announcements were made at a mosque, suggesting that our findings are indeed driven by the identity of the messenger. While these results do not rule out that a reminder effect may contribute to the magnitude of our results, they do help to illustrate that the identity of the information source plays a crucial role in generating the observed effects.

To get a better understanding of the mechanisms driving the behavioral adjustments, we consider the campaign's effects on knowledge and beliefs about the virus. While male individuals significantly adjust both labor supply and social interactions, these effects do not seem to be accompanied by changes in knowledge about the virus. Rather, the effects appear to be driven by increased concerns about the risk of transmitting the virus to others, suggesting a prosocial motive for changes in behavior. Among female individuals, on the other hand, the *phone plus loudspeaker* treatment significantly increases knowledge, but this does not translate into adjustments of behavior. This absence of an effect may be partly due to a large and significant decrease in the perceived cost of getting infected among female respondents.

Our paper makes contributions to three strands of the literature: on the relationship between religion and economic outcomes, on the role of local elites, and on the effectiveness of information interventions in the context of widespread health crises. Previous work has shown that religious institutions can influence a wide array of outcomes, ranging from the economic and social, to the political sphere. This includes aspects such as criminal behavior (Sharma, 2017; Moreno-

Medina, 2023), drinking and drug use (Gruber and Hungerman, 2008), human capital (Gruber, 2005; Becker and Woessmann, 2009), fertility and marital choices (Gruber, 2005; Bassi and Rasul, 2017), and political attitudes and participation (Basten and Betz, 2013; McClendon and Riedl, 2015; Freedman, 2020; Sperber and McClendon, 2022).[1] Most closely related, Stroebel and van Benthem (2012) finds that the appointment of a local bishop with pro-contraception attitudes in Kenya substantially increased the use of condoms among Catholic married couples. Bassi and Rasul (2017), on the other hand, illustrates how persuasive messages related to fertility during Papal speeches in Brazil increased fertility in the long run, showing that the influence of religious leaders can go both ways. We contribute to this literature by providing evidence that messages distributed by religious leaders affect health-related beliefs *and* choices beyond fertility, the health choice most closely linked to religious doctrine.

Second, our paper adds to literature on local elites and power holders in the context of low state capacity. Existing research studies the role of such elites in the implementation of local development projects (Dasgupta and Beard, 2007; Labonne and Chase, 2009; Beath et al., 2017; Casey et al., 2023), education (Reinikka and Svensson, 2004), local governance (Acemoglu et al., 2014; Sánchez de la Sierra, 2020), and with respect to land rights (Banerjee and Iyer, 2005; Goldstein and Udry, 2008). Much of this literature focuses on the potential negative effects of elites through elite capture (Bardhan, 2002; Bardhan and Mookherjee, 2006). By contrast, Sánchez de la Sierra (2020) finds that governance by local armed forces can have welfare-improving effects. Similarly, Balán et al. (2022) shows how local elites can support the state in effectively implementing property tax collection. Most closely related to our own research, Vyborny (2021) shows how religious leaders can take on a crucial role in supporting the government in the implementation of policies aimed at containing the spread of the COVID-19 virus. Her work looks at the effects of engaging with religious leaders on their likelihood to inform congregants about governmental COVID-19 rules at mosques. Our research adds to these findings by highlighting the effects of a successfully implemented information campaign by religious leaders on individuals' beliefs and behavior related to the virus.

Lastly, we contribute to a growing body of work on the determinants of the effectiveness of information campaigns as key containment strategies in the context of health hazards such as the COVID-19 virus.[2] Research has shown that trust in the source of information plays a crucial role for the successful uptake of preventive behaviors by the public at large. Empirical evidence sug-

---

[1] Other dimensions studied in previous work are income (Gruber, 2005; Bryan et al., 2020), tax morale (Torgler, 2006), donations (Condra et al., 2019; Auriol et al., 2020), values of equality and harmony (Clingingsmith et al., 2009), attitudes towards immigrants (Ben-Nun Bloom et al., 2015), as well as economic growth (Campante and Yanagizawa-Drott, 2015; Cantoni, 2015) more broadly.

[2] While we focus here on information interventions specifically related to disease containment, scholars have shown that information can have positive effects along a variety of health-related dimensions, including the use of water purification (Ashraf et al., 2013), responses to maternal health risks (Ashraf et al., 2023), and risky sexual behavior (Dupas et al., 2018).

gests that local influential figures may serve as key trustworthy disseminators of public health information—such as immunization reminders (Banerjee et al., 2019, 2021; Alatas et al., 2024) and COVID-19 preventive measures (Banerjee et al., 2020; Solís Arce et al., 2021; Vyborny, 2021)—to their communities. Religious leaders, who have substantial influence over adherents and credibility within their communities, may be particularly valuable partners in the implementation of public health efforts. During the 2014–2016 Ebola pandemic, religious leaders in Sierra Leone advocated for protective health practices such as hand-washing and safe burials (Greyling et al., 2016). Vyborny (2021) shows that one-to-one engagement with religious leaders motivates them to promote government policies to contain the spread of COVID-19 during religious gatherings.[3] On the outcome side, Abaluck et al. (2022) shows that a combined mask distribution and promotion intervention featuring a religious component significantly increased mask use and reduced symptomatic COVID-19 infections. The study does, however, not look at the impact of communication by religious leaders separately. Relatedly, Armand et al. (2022) find that signalling social proximity—such as religious concordance—between the source and receiver of information helps dispel myths and misinformation in the context of the COVID-19 outbreak, thus contributing to the effectiveness of information campaigns. To the best of our knowledge, our study is the first to investigate the effects of public health information spread by influential religious figures on preventive behavior. In particular, we show that receiving messages by religious leaders significantly decreased labor supply and social interactions during the COVID-19 pandemic.

2.2   Background and Experimental Design

Our study was implemented during the COVID-19 pandemic in two provinces of Pakistan, Punjab and Sindh. In this setting, religion plays a major role. 97% of the population are Muslim and the vast majority of individuals (90%) considers religion to be very important in their lives (Pakistan Bureau of Statistics, 2017b; Haerpfer et al., 2022). Not surprisingly, a high level of trust in religious institutions and authorities is almost universal (97%) (Haerpfer et al., 2022). At the same time, trust in secular institutions is much lower. According to the latest wave of the World Value Survey in 2020, 53% of the respondents have little to no confidence in political parties (Haerpfer et al., 2022). Potentially even more important in the context of a global health crisis, only 39% of the interviewed individuals report having at least some confidence in the

---

[3]   While religious leaders can have substantial positive effects on the adoption of health-beneficial behavior, their role hinges on their support for the respective health measures. If they are skeptical of certain campaigns, their widespread influence has the potential to provoke behavior with adverse health effects. For instance, Jegede (2007) illustrates how conspiracy theories spread by religious leaders in Northern Nigeria led to a boycott of the Polio vaccination campaign. Similarly, Martinez-Bravo and Stegmann (2022) shows that an anti-vaccine propaganda campaign by the Taliban in Pakistan led to significant declines in immunization rates.

World Health Organization (WHO), and the vast majority (94%) agrees with the statement that, whenever science and religion conflict, religion is always right (Haerpfer et al., 2022). These facts emphasize the crucial role of religious institutions in shaping individuals' attitudes and behavior and highlight the importance of their support in effectively implementing containment strategies.

In Pakistan, the first COVID-19 cases were confirmed in March 2020, with local transmission rapidly spreading in densely populated cities. While the initial responses of Pakistan's provincial governments varied, a nationwide lockdown was imposed starting March 24, 2020, and lifted in phases beginning May 9, 2020.[4] In cooperation with the National Rural Support Program (NRSP), a local NGO, we implemented a COVID-19 awareness campaign in the time period between September and November 2020, after the first peak of cases.[5]

NRSP currently works with more than 3.5 million poor, primarily rural, households. To implement this project, NRSP provided us with all available phone numbers of their beneficiaries in Punjab and Sindh—more than 50,000 in total. We grouped these individuals geographically by revenue village, with the goal of interviewing around 15 randomly selected beneficiaries per village at baseline.[6] During the baseline survey, enumerators collected information on the socio-economic characteristics and health status of the respondents and their household members, their knowledge and perceptions related to COVID-19, as well as any preventive measures taken by the respondents or their household members. This data was collected between August and October, 2020.

We randomized each revenue village into one of three treatment conditions.[7] Out of a total of 904 villages, one third was randomly assigned to a pure control group, in which individuals received no awareness intervention at all. Three quarters of the remaining villages were allocated to receive awareness messages transmitted via phone calls only, and one quarter was allocated to receive COVID-19 loudspeaker announcements by religious leaders in addition to awareness phone calls.[8] The randomization was successful in generating a well-balanced sam-

---

[4] The authorities developed the National Action Plan for COVID-19 that provided guidelines for priority testing, social distancing, quarantine facilities, and standard operating procedures for events such as Ramadan, Eid, gatherings, ceremonies, and marriages. Further guidelines were also specified for educational institutions, tourism services, and air transportation.

[5] See https://coronavirus.jhu.edu/region/pakistan for an overview of COVID-19 cases in Pakistan over time.

[6] A revenue village is the smallest unit in the administrative hierarchy of Pakistan, with a median population of 2,259 inhabitants in the study area (Pakistan Bureau of Statistics, 2017a). We excluded all revenue villages for which less than 20 beneficiary phone numbers were available prior to the baseline survey. For the remaining revenue villages, the response rate to our baseline call lies at 39.5%. Since we do not have any pre-baseline information on the beneficiaries, we cannot, however, investigate whether selection into responding to our baseline call is correlated with respondent characteristics.

[7] Prior to randomization, we excluded all individuals who did not consent to participating in any further interviews as well as individuals with disproportionally much missing information at baseline (above the 99th percentile).

[8] Due to budget as well as time constraints among the implementing staff, we could allocate only a smaller

ple with respect to the main baseline characteristics, with minor imbalances in terms of respondents' experiences with individuals being treated badly because of a COVID-19 infection, their beliefs about traditional healers being able to treat COVID-19, and their perceptions about the cost associated with contracting the COVID-19 virus. We summarize the experimental design in Table 2.1 and show randomization balance in Appendix Table 2.B.1.

Table 2.1: Experimental Design

	Randomization		Endline		Final sample	
	Villages	Individuals	Villages	Individuals	Villages	Individuals
	(1)	(2)	(3)	(4)	(5)	(6)
Phone	452	5,399	448	3,399	445	3,183
Phone + loudspeaker	150	1,896	148	1,187	147	1,112
Control	302	3,640	297	2,344	294	2,188
Total	904	10,935	893	6,930	886	6,483

Notes: Columns (1) and (2) show the number of villages and individuals randomly assigned to the three treatment groups after the baseline survey. Columns (3) and (4) show the number of villages and individuals who responded to the endline survey. Columns (5) and (6) show the number of villages and individuals in the final sample, i.e., all individuals for whom we have information on the three main outcome indices (labor supply, social interactions, and preventive measures) at both baseline and endline. Changes in the amount of villages across stages reflect individual level attrition in villages with very few respondents to begin with. On average, villages which are not included in the final sample had four baseline respondents, as compared to thirteen in villages included in the final sample.

Our project included two additional experimental variations that we do not study independently in this paper. First, individuals in both the *phone* and the *phone plus loudspeaker* group were cross-randomized on the individual level to receive one of five different types of awareness calls. Whereas everyone received basic information about the virus, four out of five groups additionally received information with respect to either the severity of the virus, the risk of infection, the risk of infecting others, or the cost related to contracting the virus. We abstract from this variation and consider only the average effect of having received any type of awareness message.[9] Second, in a random third of the *phone plus loudspeaker* villages, NRSP employees additionally distributed COVID-19 information posters in public spaces. We control for assignment to the poster treatment throughout the regressions presented in this paper.

Awareness Phone Calls   Our awareness calls consisted of (1) an introduction to what COVID-19 is and how it is transmitted, (2) a description of the main risk groups and symptoms of COVID-19, (3) information on recommended preventive behavior including hygiene practices, social distancing, and wearing of a mask, and (4) recommendations on how to react in case the

---

fraction of villages to the combined phone and loudspeaker treatment arm.

[9]   We do not control for the specific treatment message in our main specification, but discuss the robustness of our results to the inclusion of this indicator in Section 2.4.3.

respondent suspects that they or a household member are infected with the virus. The content was developed by the research team in Germany, and contextualized by local partners based in Pakistan. The campaign provided information circulated by reliable sources (e.g., the World Health Organization (WHO), University College London, John Hopkins Medicine, and the Center for Disease Control and Prevention), used simplified terms, and was available in Urdu and Sindhi, the two local languages spoken in the study area. The complete message scripts can be found in 2.A. The awareness phone calls lasted 15 to 20 minutes and were conducted by trained employees of our local partner NRSP. To keep the respondents' attention, the call was designed in an interactive way, stopping and asking questions at times, and repeating information where necessary. Our monitoring data shows that we reached approximately 74% of all individuals assigned to receive an awareness call.

Awareness Loudspeaker Announcements   Mosque loudspeakers are conventionally used in the study area to disseminate information to the public. This way of communication allows spreading messages widely without involving personal interaction. We mobilized Imams to make COVID-19 related announcements via mosque loudspeakers in 147 treated villages. Our implementing partner, NRSP, engaged its social mobilization staff to identify two community activists per village to interact with the Imams and convince them to make the announcements.[10] The community activists explained the content of the message to the religious leaders and, in cooperation with other community members, ensured its delivery via loudspeakers. The messages can be understood as a brief summary of the content transmitted during the phone calls.[11]   The announcements, conducted on average twice per day on four days a week, were scheduled between noon and afternoon, especially on Jummah (the Friday Prayer), an important day of prayer for Muslims.

According to NRSP's monitoring statistics, the announcements were completed in every assigned village. As reported in Table 2.2, an average of 147 announcements were made per village in the course of the implementation period, which lasted approximately 22 days. On average, 25 announcements were delivered per week in 7 mosques per village.[12] As visible from the summary statistics, there is substantial variation in the amount of announcements made

---

[10] Community activists are members of the community who had worked with NRSP in the past and provided support for implementing various projects before our intervention.

[11] See 2.A for the complete loudspeaker script.

[12] We also ask individuals about both awareness calls and loudspeaker announcements they received at end-line. We show descriptive statistics on self-reported take-up of such interventions—either by NRSP or another organization—by treatment group in Appendix Table 2.B.2. While there is no significant difference across experimental groups with respect to the likelihood or recalling a COVID-19 related call, individuals assigned to the *phone* and *phone plus loudspeaker* treatments are significantly more likely to report having received information about the symptoms and transmission of, and preventive measures against, COVID-19 during such a call. Importantly, individuals in the *phone plus loudspeaker* treatment are also significantly more likely to report having received information via loudspeaker announcements made at a mosque.

Table 2.2: Descriptive Statistics on the Implementation of Loudspeaker Announcements

	Mean	SD	Min.	Max
	(1)	(2)	(3)	(4)
# of announcements made per village	146.96	114.15	12	672
# of mosques/ places covered with announcements per village	6.58	4.89	1	28
# of days announcements were conducted in a village	21.66	4.02	11	26
# of villages covered with announcement per week	138.04	15.17	111	150
# of announcements made per village per week	24.88	22.19	0	112

Notes: The table reports summary statistics from the monitoring data on the implementation of loudspeaker announcements.

across villages. This likely reflects differences in the NGO's ability to engage with local religious leaders and motivate them to support the awareness campaign, and suggests that treatment intensity differs across villages assigned to the *phone plus loudspeaker* treatment. While we are unable to causally identify the effect of this variation due to its endogeneity, we will present heterogeneity analyses by treatment intensity to get a better idea of the correlation between treatment intensity and our outcomes of interest.

Following the implementation of our awareness campaign, we conducted an endline survey with individuals in both the treatment and control groups. The survey took place between December 2020 and January 2021, on average two months after the awareness call for treated individuals. We used this survey to elicit information about the main outcomes of interest, namely labor supply, social interactions, and preventive measures, as well as perceptions related to the COVID-19 virus. For consistency, in our final analysis, we include only those individuals for whom we have information on the three main behaviors related to the prevention of a COVID-19 infection at both baseline and endline.[13]

2.3   Descriptive Statistics and Measurement

Descriptive Statistics   Table 2.3 presents summary statistics for a wide range of baseline characteristics and baseline levels of COVID-19 related knowledge, perceptions, and behavior. The average respondent is around 38 years old, male, completed four years of education (4th grade), and either took on daily-wage jobs—e.g., construction work, factory work, or street vending—

---

[13] We present balance results for the final sample in Appendix Table 2.B.3. The final sample is well-balanced across the main characteristics with few exceptions, in particular age, the perceived cost of getting infected, and individuals' knowledge with respect to traditional healers' ability of curing the COVID-19 virus.

Table 2.3: Baseline Characteristics

	Mean (1)	SD (2)	Min. (3)	Max. (4)	# (5)
Panel A: Respondent Characteristics					
Age	38.29	11.21	18	80	6,463
Male respondent	0.52	0.50	0	1	6,483
Highest level of education completed	4.27	4.85	0	15	6,474
Daily wage laborer or unemployed (BL)	0.83	0.37	0	1	6,483
# of household members (w)	8.44	4.22	1	30	6,432
Income in the last 7 days (w)	841.04	1267.40	0	4000	6,243
Panel B: COVID-19-related Knowledge					
Knows symptoms of corona virus	0.86	0.35	0	1	6,117
Frequently and thoroughly wash hands with soap	0.79	0.40	0	1	6,119
Wear a face mask	0.58	0.49	0	1	6,119
Maintain two meters distance from people	0.50	0.50	0	1	6,119
Panel C: COVID-19 Perceptions					
Thinks one would die or recover with severe health damages if infected	0.51	0.50	0	1	4,671
Has moderate or high concerns of getting infected	0.29	0.45	0	1	5,998
Has moderate or high concerns of infecting other household members	0.41	0.49	0	1	5,883
Panel D: COVID-19-related Behavior					
Preventive measures					
Frequently and thoroughly washes hands with soap	0.76	0.42	0	1	5,645
Not touching any objects used by other people before washing them with soap	0.12	0.33	0	1	5,659
Maintains two meters distance from people	0.40	0.49	0	1	5,653
Wears a face mask	0.45	0.50	0	1	5,305
Social interactions					
Left village (last 7 days)	0.23	0.42	0	1	6,476
Went to the market (last 7 days) (resp. or household member)	0.71	0.45	0	1	6,446
Went to the mosque, church or mandir (last 7 days)	0.56	0.50	0	1	6,197
Labor supply					
Worked outside home (last 7 days)	0.38	0.49	0	1	6,483
# of days worked outside home (last 7 days, cond.)	4.97	2.26	0	7	2,388
# of hours worked outside home yesterday (cond.)	6.77	3.58	0	14	2,424

Notes: The table reports summary statistics for the final sample used in the analysis. Missing observations are due to respondents indicating that they do not know the answer to a question or do not wish to reply to a question (overall sample size: $N = 6,483$).

or did not work in the seven days preceding the baseline survey.[14] The average household has eight members, and average reported income in the seven days prior to the interview amounts to 841 Pakistan Rupees, then approximately 5.1 US dollars.[15]

At baseline, 86% of respondents reported to know the symptoms of COVID-19. The most widely known preventive measures included washing hands with soap (79%), wearing face masks (58%), and maintaining two meters distance from others (50%). Half of respondents thought they would die or recover with severe health damage if infected. Yet, only 29% of respondents had moderate to high concerns about getting infected, and 41% about infecting others. This suggests that individuals likely overestimated the severity of the COVID-19 virus, but underestimated the probabilities of getting infected and infecting others at baseline.

The most widely practiced preventive measure was frequent and thorough hand-washing (76%), followed by wearing a face mask (45%) and maintaining two meters distance from others (40%). On average, 23% of respondents left their village in the seven days prior to the interview. 71% went to the market (respondents themselves or another household member), and 56% went to a religious institution. Finally, 38% of respondents worked outside home in the seven days prior to the survey, for on average five days.

Table 2.B.4 in the Appendix reports gender-differentiated baseline summary statistics revealing that at baseline, a considerably higher share of male, as compared to female, respondents had visited a mosque, church, or mandir in the seven days prior to the interview (75% versus 35%). Similarly, the labor behavior at baseline differs substantially between male and female respondents, with male individuals having worked substantially more frequently outside home than female individuals (50% versus 26%).

Outcome Measurement   We aim to test the awareness campaign's impact on individuals' perceptions and knowledge about COVID-19, as well as on their related behavior. In particular, we are interested in potential effects on labor supply, social interactions, and the use of preventive measures. To look at these topics, we make use of self-reported data from our endline survey. For each outcome of interest, we construct an index that combines several related variables. Following Kling et al. (2007), we build these indices in two steps. First, we standardize each variable by subtracting its control group mean and dividing by its control group standard deviation. We then compute a simple average of all standardized components of an index by

---

[14] Based on our survey questions, we are unable to distinguish between individuals who didn't work in the seven days prior to the baseline interview, but are otherwise active in the labor market—i.e., individuals who would likely be considered daily wage laborers—and inactive individuals. Non-daily wage laborers were employed as/in: skilled labor, personal business, self-cultivator/own farm, cultivation on contract, cultivation on partnership/share cropper, family helper, employer/business, livestock only.

[15] The conversion is based on the average exchange rate between August and October 2020—the three months in which baseline interviews were conducted—taken from https://www.exchangerates.org.uk.

summing the respective values and dividing by the number of components of an index.[16]

We construct three indices to investigate the effects of the treatment on individuals' behavior related to the virus. The *labor supply index* contains information about the frequency with which a respondent worked outside home in the week prior to the interview. The *social interactions index* combines information on visits received by the respondent, their participation in social and religious gatherings, market visits, and their handling of social interactions more generally. The *preventive measures index* includes information on the use of masks, hand-washing, and distancing of at least two meters as a means to reduce the risk of infection. Each of these indices increases in the extent to which an individual takes measures to contain the spread of COVID-19, i.e., with decreasing labor supply and social interactions, and increasing preventive measures.

To explore the effects on *knowledge*, we construct an index that measures an individual's level of information about the symptoms and transmission channels of COVID-19, as well as potential measures to prevent an infection. This index is larger the more informed an individual is. Finally, we want to explore changes in beliefs related to the virus. We use two indicator variables to capture whether an individual has moderate or high concerns about *getting infected* or *transmitting* COVID-19, respectively. To look at the *perceived cost of an infection*, we construct an index that combines information on the expected time necessary for recovery, as well as the estimated cost of treatment and foregone income due to time spent without working. Finally, we look at an indicator for the *perceived severity* of the virus, capturing whether individuals believe that one would die or recover from an infection with severe health damages, as compared to recovering fully or with less severe consequences. The three dummy variables and the index increase in the perceived risk of infection and transmission, the perceived cost of an infection, and the perceived severity of the virus. A complete list of the variables used in the construction of all indices can be found in Appendix Table 2.B.5.[17]

Estimation Strategy   To analyze the effects of our awareness campaign, we estimate the following equation using OLS:

$$Y_{iv} = \beta_0 + \beta^{ph} T_{iv}^{ph} + \beta^{ph\&ls} T_{iv}^{ph\&ls} + \eta X_{iv0} + \zeta + \epsilon_{iv}, \tag{2.1}$$

---

[16] Prior to standardization, all components of an index are coded such that a higher value can be interpreted in a consistent way. For example, in the labor index, all components are recoded such that a higher number reflects lower labor supply.

[17] Unless explicitly indicated, we include the single variables pre-specified in the pre-analysis plan (PAP). Wherever we do not have multiple outcomes of interest on one topic, we use the outcome variables without standardizing them.

where i indexes the phone number called to reach the respective individual and v indexes the revenue village.[18] T_{iv}^{ph} is an indicator for villages that were assigned to receive the awareness campaign via phone calls only, whereas $T_{iv}^{ph\&ls}$ is an indicator for villages assigned to the awareness campaign via phone calls and loudspeaker announcements. Hence, β^{ph} and $\beta^{ph\&ls}$ estimate the intention to treat (ITT) effects of the two treatment arms on the respective outcome of interest (Y_{iv}). X_{iv0} corresponds to a matrix of covariates that includes the baseline values of those village- and individual-level variables that were used in the randomization procedure, an indicator for whether a village was assigned to the poster treatment, as well as a variable capturing the baseline level of Y_{iv}. ζ represents enumerator fixed effects. Finally, standard errors are clustered at the village level.

2.4 Main Results

2.4.1 Treatment Effects on Labor Supply, Social Interactions, and Preventive Measures

In this section, we consider whether the two treatments were successful in adjusting individuals' behavior to reflect a more cautious dealing with the pandemic. Table 2.4 summarizes ITT estimates from equation 2.1 for the three indices of interest—labor supply, social interactions, and preventive measures. We present our estimates for the overall sample (Columns 1, 4, 7), as well as for male (Columns 2, 5, 8) and female respondents, separately (Columns 3, 6, 9). In this and all of the following tables, the first row presents the effects of the awareness campaign conducted via phone calls alone (*phone* treatment), while the second row shows the effects of the combined intervention via phone calls and loudspeaker announcements (*phone plus loudspeaker* treatment). p-values for the significance of the difference between the two treatment arms are shown below. Similarly, we show p-values for the average effect of being treated with either the phone or the phone + loudspeaker treatment.

   Our results show that the awareness campaign conducted via phone calls alone did not have a statistically significant effect on any of the indices. What appears to have made a difference, though, is the additional assignment to loudspeaker announcements by Imams. In particular, the *phone plus loudspeaker* treatment led to a significant downward adjustment in labor supply by 0.06 standard deviations in the overall sample. The difference in effect sizes between the

---

[18] We attempt to reach the same individuals over time and implement several checks to identify the respondent who gave consent to the research study. In practice, we are only certain that the same number was called. In some instances it may thus have happened that different individuals picked up the phone during the baseline, the awareness, and the endline call. To simplify notation, we will refer to individual-level observations while we actually capture contact-level observations attempting to verify whether we speak to the same individual over time.

Table 2.4: Effects on Labor Supply, Social Interactions, and Preventive Measures

	Less Labor Supply			Less Social Interactions			More Preventive Measures		
	All (1)	Male (2)	Female (3)	All (4)	Male (5)	Female (6)	All (7)	Male (8)	Female (9)
Phone	0.008	0.010	-0.011	0.008	0.007	-0.000	0.014	0.003	0.016
	(0.022)	(0.028)	(0.030)	(0.012)	(0.015)	(0.018)	(0.014)	(0.017)	(0.022)
Phone + loudsp.	0.060**	0.089***	0.049	0.014	0.044*	-0.007	0.014	0.011	0.010
	(0.031)	(0.044)	(0.044)	(0.017)	(0.024)	(0.025)	(0.024)	(0.030)	(0.035)
p-val phone=phone+loudsp.	0.077	0.064	0.133	0.709	0.128	0.760	0.989	0.785	0.850
p-val any treatment	0.395	0.382	0.975	0.437	0.362	0.926	0.295	0.781	0.489
Obs.	6,483	3,370	3,113	6,483	3,370	3,113	6,483	3,370	3,113
Villages	888	665	735	888	665	735	888	665	735
R²	0.360	0.372	0.395	0.408	0.367	0.432	0.425	0.442	0.439
Control mean	-0.000	-0.205	0.224	0.000	-0.139	0.152	-0.000	-0.007	0.007
SD	0.946	0.940	0.901	0.548	0.475	0.582	0.621	0.590	0.653

Notes: The table shows treatment effects on summary indices of labor supply, social interactions, and preventive measures. Results are reported for the over-all sample, and the male and female subsamples, respectively. Significance levels are indicated by $*\ p < 0.10$, $**\ p < 0.05$, $***\ p < 0.01$.

phone and the *phone plus loudspeaker* treatment is statistically significant.

The effect on labor supply is driven by male individuals, for whom exposure to the combined treatment led to a decrease of 0.09 standard deviations in the labor supply index. The fact that we observe adjustments in labor among men, but not women, may well be driven by their much higher baseline participation in the labor market. This is true not only in our study area, but also more broadly in the context of Pakistan.[19]

We also find evidence, albeit somewhat weaker, for a downward adjustment in social interactions among male individuals by around 0.04 standard deviations. We detect no significant effects—neither overall, nor gender-differentiated—on the preventive measures index. It is worth mentioning, however, that hand washing, mask wearing, and distancing have increased throughout the study period in both the control and the two treatment groups (see Appendix Table 2.B.6). This increase may be driven by the simultaneous implementation of other COVID-19 information campaigns in the study area, specifically targeting the most common preventive measures, and thus may explain why our campaign has no additional effect on preventive measures among individuals assigned to either of the two treatment arms.

Effects on the Components of the Labor Supply Index   To better understand the observed adjustments in labor supply, we look at all components of the index separately. Figure 2.1 displays the effects of the *phone plus loudspeaker* treatment on the three components of the labor supply index for the overall sample and the male subsample, respectively.[20] We find significant reductions of labor supply at both the extensive and intensive margins, driven entirely by male individuals. In the overall sample, the share of individuals who reported working outside home in the seven days prior to the endline interview is 3.6 percentage points (7.4%) lower in the *phone plus loudspeaker* treatment group than in the control group (control group mean: 48.6%). When zooming in on male individuals, the magnitude of the reduction raises to 4.8 percentage points, a reduction of 8.0% as compared to the control group mean of 60%.

The assignment to phone calls and loudspeaker announcements by Imams also led to a significant drop in the number of days worked outside home. The decline amounts to about one-quarter of a workday in the overall sample (control group mean: 2.5) and almost half a workday in the male subsample (control group mean: 3.2). While we observe a negative coefficient for the number of hours worked outside home on the day prior to the interview in the *phone plus*

---

[19] In our sample, the share of male individuals who reported working outside home in the 7 days prior to the baseline interview amounts to 50%. For female individuals, it is as low as 26%. This is in line with the official statistics of the International Labour Organization, which report a female labor force participation rate of 33.6% in rural Pakistan (International Labor Organization, 2024). Tanaka and Muzones (2016) documents as a major self-reported reason for women not working outside home in Pakistan their family members' objection to the latter.

[20] Figure 2.B.1 in the Appendix displays the effects for the female subsample.

loudspeaker treatment group, this effect is not statistically significant. We further do not find any effects on either of the components of the labor supply index for female respondents (see Figure 2.B.1 in the Appendix).

We also test whether a reduction in the likelihood of working *outside* home is accompanied by an increase in the likelihood of working *from* home, but do not find evidence for a compensation effect. The majority of working individuals in our sample pursue daily-wage jobs, such as construction or factory work, which usually require presence at the workplace. At baseline, the share of individuals working from home is extremely low at only 4.4%, and even lower among men (2.7%). This indicates that the poor, mostly unskilled, individuals in our sample are largely unable to do their jobs from home and as such to compensate for a reduction in work outside home.[21]

Figure 2.1: Effects on Single Outcomes of the Labor Supply Index: Overall Sample and Male Subsample
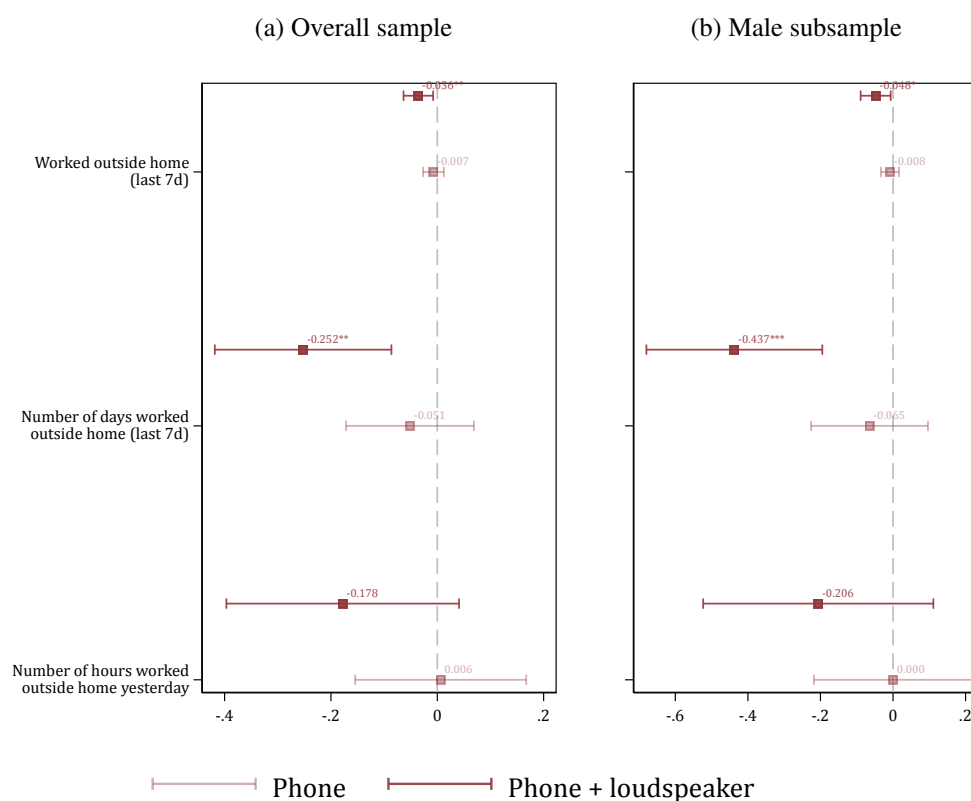
Notes: The graph shows treatment effects on the components of the labor supply index for the overall sample (left) and the male subsample (right). Point estimates are shown with 10% confidence intervals. Significance levels are indicated by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Overall, these results suggest that the impact of our intervention on individuals' behavior

---

[21] In absence of a compensation effect, the decrease in overall labor supply leads to a significant decrease in weekly income of around 198.7 rupees among individuals assigned to the *phone plus loudspeaker* treatment, a reduction of 15.3% as compared to the control group mean.

is driven by the additional effort in villages in which religious leaders were mobilized to endorse the awareness campaign. In particular, as a result of involving religious leaders, male individuals responded by significantly reducing labor supply. The magnitude of this reduction is striking given that the loudspeaker messages delivered by the Imams did not specifically address labor supply, but rather focused on preventive measures (hand washing, use of face masks, and social distancing). Leveraging on the authority of religious leaders may have created social pressure for compliance with COVID-19 safety measures more broadly, leading individuals to temporarily cut back on their activities.

Given the substantial effect on labor supply, it is natural to ask whether the reduction in working hours among individuals in the *phone plus loudspeaker* group translated into a reduction in infection rates. While we ask individuals to report their own, as well as their household members', health status, we observe very little incidences of sickness—likely due to underreporting. Anecdotal evidence from conversations with our implementing partner suggests that individuals in our study setting may have been uncomfortable reporting sickness out of fear that information might be shared with the government. This highlights once more the importance of trust in governmental and non-governmental actors trying to contain the spread of a disease. The small number of reports of sickness within our sample leaves us underpowered to identify any potential effects on self-reported infections.

2.4.2 Treatment Effects on Knowledge and Beliefs

To get a better understanding of the underlying drivers of the observed behavioral change, we turn to the campaign's effects on knowledge and beliefs related to the COVID-19 virus. Table 2.5 illustrates the effects of the two treatment arms on an index that combines indicators of COVID-19 knowledge. Overall, only the *phone plus loudspeaker* treatment led to significant increases in knowledge. However, the coefficient is not significantly different from that of the *phone* treatment.

Strikingly, the effects on knowledge are largely driven by female respondents, suggesting that the changes in labor activities among male respondents cannot be explained by changes in knowledge about the virus. Women, on the other hand, have significant learning effects, but these effects do not translate into changes in behavior. This may be partly due to the type of information that women retain from the awareness campaign.[22] On average, female respondents who received the *phone plus loudspeaker* treatment are 3.2 percentage points more likely to report that they know the symptoms of the COVID-19 virus. This subjective measure is in line with a significant increase of 0.27 correctly reported symptoms. Similarly, women in the *phone*

---

[22] Figure 2.B.2 in the Appendix illustrates the effects of the treatment on the components of the knowledge index for the overall sample as well as female respondents.

Table 2.5: Effects on Knowledge

	More Knowledge		
	All	Male	Female
	(1)	(2)	(3)
Phone	0.020	0.003	0.033*
	(0.013)	(0.019)	(0.018)
Phone + loudsp.	0.042**	0.042	0.053*
	(0.019)	(0.027)	(0.027)
p-val phone=phone+loudsp.	0.248	0.139	0.440
p-val any treatment	0.050	0.571	0.033
Obs.	5,741	2,975	2,766
Villages	879	644	708
R^2	0.384	0.374	0.413
Control mean	0.007	0.061	-0.052
SD	0.526	0.499	0.549

Notes: The table shows treatment effects on the knowledge index. Results are reported for the overall sample, and the male and female subsamples, respectively. Significance levels are indicated by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

group name on average 0.12 more correct symptoms. Finally, female respondents in the *phone* and the *phone plus loudspeaker* group are substantially more likely to know that the COVID-19 virus can be contracted by touching contaminated surfaces (3.3pp and 4.0pp, respectively).

Beyond purely factual knowledge, individuals may also update their subjective beliefs about the virus in response to the treatment. Thinking about the trade-off between the cost and benefit of adjusting behavior, four beliefs about the virus may be particularly relevant determinants of behavioral changes: perceptions about the probability of infection, the transmission probability, the costs related to getting infected, and the severity of the disease. In Table 2.6, we report estimates on the four aforementioned perceptions for the overall sample, as well as the male and female subsamples, respectively. While perceptions about the cost of getting infected combine information on the perceived cost of treatment and forgone income based on the expected duration of the infection into an index, all other outcomes are coded as indicator variables.

Our findings reveal that both treatments had significant effects on the perceived risk of transmitting the virus to others. Individuals assigned to the *phone* (*phone plus loudspeaker*) treatment are on average 2.1 (3.2) percentage points more likely to have moderate or high concerns about infecting their household members with the virus. While the coefficient is larger among individuals in villages in which the awareness campaign was endorsed by Imams, there is no significant difference between the two treatment effects. Looking at the male and female subsample separately, the effect remains significant only among men who received the *phone plus loudspeaker* treatment, who are about 4.3 percentage points more likely to have concerns about

Table 2.6: Effects on Beliefs

	Higher Perceived Risk of Getting Infected			Higher Perceived Risk of Transmission			Higher Perceived Cost			Higher Perceived Severity		
	All (1)	Male (2)	Female (3)	All (4)	Male (5)	Female (6)	All (7)	Male (8)	Female (9)	All (10)	Male (11)	Female (12)
Phone	0.010	0.012	0.001	0.021**	0.015	0.017	0.025	0.010	0.032	-0.009	-0.014	-0.011
	(0.010)	(0.013)	(0.016)	(0.009)	(0.011)	(0.015)	(0.029)	(0.041)	(0.044)	(0.013)	(0.020)	(0.019)
Phone + loudsp.	0.011	0.022	-0.007	0.032**	0.043**	0.010	-0.044	0.020	-0.114***	-0.018	-0.006	-0.016
	(0.016)	(0.021)	(0.024)	(0.015)	(0.020)	(0.022)	(0.038)	(0.066)	(0.042)	(0.019)	(0.026)	(0.030)
p-val phone=phone+loudsp.	0.944	0.613	0.744	0.417	0.164	0.723	0.082	0.876	0.004	0.646	0.763	0.857
p-val any treatment	0.298	0.283	0.957	0.008	0.054	0.285	0.638	0.764	0.930	0.350	0.484	0.518
Obs.	5,555	2,882	2,673	5,441	2,844	2,597	3,391	1,841	1,550	3,805	1,878	1,927
Villages	876	639	699	875	635	691	764	533	547	806	553	608
R^2	0.541	0.587	0.507	0.640	0.712	0.576	0.417	0.427	0.435	0.535	0.572	0.540
Control mean	0.375	0.349	0.403	0.412	0.378	0.450	0.008	0.051	-0.040	0.531	0.539	0.524
SD	0.484	0.477	0.491	0.492	0.485	0.498	0.850	0.906	0.780	0.499	0.499	0.500

Notes: The table shows treatment effects on the perceived risk of getting infected, the perceived risk of transmitting the virus, the perceived cost of getting infected, and the perceived severity of an infection. Results are reported for the overall sample, and the male and female subsamples, respectively. Significance levels are indicated by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

the transmission of the virus. This indicates that an increase in concerns about the health of others, rather than a concern for oneself, may have motivated the observed behavioral changes. These results emphasize once more the critical role of religion in motivating altruistic beliefs and behavior.

Finally, we observe a significant decrease in the perceived cost of getting infected with the COVID-19 virus among women exposed to the *phone plus loudspeaker* treatment. While this result may reflect a realistic adjustment based on better knowledge of the characteristics of the virus, a lack of data on the actual cost of getting infected complicates the interpretation of the effect. Given that the monetary repercussions of a COVID-19 infection might be particularly important in the context of an extremely poor sample, the significant decrease in concerns about the latter may, however, add important context to why we observe no behavioral response among women in the *phone plus loudspeaker* group.

2.4.3   Robustness of Results

We present results on the robustness of our main analyses in Appendix Tables 2.B.7, 2.B.8, and 2.B.9. Our findings are robust to the exclusion of individual and village level characteristics as control variables, as well as to including only those individual level characteristics that are unbalanced across treatment arms. We also show additional specifications controlling for the specific treatment message received during the awareness call. While the results of these regressions do not directly serve as robustness check for the main findings, they provide interesting additional insights. Importantly, our main results capture the effects of any type of awareness call, as well as the combination of an awareness call and a loudspeaker announcement made by the Imam, independently of the specific awareness message. By controlling for the specific message, the additional specifications show the effects of the basic awareness call, either independently or in combination with loudspeaker announcements.

The coefficients of these regressions are similar to the ones presented in the main part of the paper, suggesting that the average effects of the two treatment arms are comparable to the effects of the two treatments among individuals who received only the basic awareness message. However, while controlling for the treatment message reduces the effect size on labor supply among male respondents, the effect becomes significant among female respondents. This indicates that the additional treatment messages related to the perceived risks of infection and transmission, the perceived cost, and the perceived severity of the virus interact differently with the loudspeaker announcements depending on the gender of the respondent.

2.5   Heterogeneity of Treatment Effects

In what follows, we consider the heterogeneity of the campaign's treatment effects on labor supply with respect to two variables of interest: employment status and religiosity. As adjustments along the behavioral indices are entirely driven by male individuals, we focus our analysis on the male subsample. Results on both the overall, and the female subsample are presented in the Appendix.

Labor Supply by Baseline Employment Status   Given the economically significant effect on labor activity and the potential effects that such reductions can have on income, it is important to understand who reduces working hours in response to the treatment. This can provide valuable insights into the mechanisms driving the effects on the labor index. To look at the effects of the treatment by employment type, we consider (1) men who did not work in the seven days prior to our baseline survey, (2) men who worked as daily wage laborers—e.g., in construction, factory work, or street vending—, and (3) men who worked in other types of jobs—e.g., in personal businesses, on their own farms, or as skilled laborers. The first category contains both long-term unemployed individuals and (daily wage) laborers who did not work in the week before the interview, but are otherwise active in the labor market. Based on our survey data, we cannot perfectly distinguish these two types of individuals. Yet, in the control group, 58% of those men who did not work at baseline indicate having worked in the seven days prior to the endline survey.[23] Thus, men actively seeking work seem to make up the majority of individuals in the first category. This also explains why we would expect any adjustments in labor supply for individuals not having worked at baseline.

Table 2.7 presents the results on the effects of being treated by employment status. As compared to the control group, the *phone plus loudspeaker* treatment led to a downward adjustment by 0.20 standard deviations among male daily wage laborers. The coefficient on the *phone plus loudspeaker* treatment is smaller and insignificant for individuals who did not work in the seven days prior, and negative, but statistically insignificant, for male individuals in other types of employment.[24] These findings suggest that individuals in more dire working conditions decreased working hours in response to the *phone plus loudspeaker* treatment, while those working, e.g., as skilled laborers, or as independent business men or farmers, did not. Given that daily-wage laborers are likely to more heavily rely on their day-to-day earnings than individuals with more permanent jobs, one would probably expect them to be more reluctant in reducing labor.[25] This

---

[23] This number is somewhat lower at 47% for the overall sample.

[24] Table 2.B.10 in the Appendix shows the same pattern of results for the overall sample. However, the effects are attenuated, as there is no reaction to the treatment among female individuals in either of the three groups.

[25] In line with this assumption, weekly income at baseline is significantly lower for daily wage laborers than for individuals with less precarious employment conditions.

Table 2.7: Effects on Labor Supply by Employment Status: Male Subsample

| | Less Labor Supply | | |
	No Work at BL (1)	Daily Wage Laborer (2)	Other Employment (3)
Phone	0.035	-0.037	-0.019
	(0.043)	(0.056)	(0.072)
Phone + loudsp.	0.093	0.196**	-0.012
	(0.058)	(0.099)	(0.107)
p-val phone=phone+loudsp.	0.326	0.011	0.950
p-val any treatment	0.253	0.917	0.802
Obs.	1,609	1,010	751
Villages	516	385	324
R^2	0.364	0.422	0.529
Control mean	-0.129	-0.289	-0.286
SD	0.939	0.909	0.973

Notes: The table shows treatment effects on the summary index of labor supply for the male subsample. The results are reported for individuals who did not work in the seven days prior to the baseline interview (1), individuals who worked as daily wage laborers (2), and individuals in other type of employments (3). Significance levels are indicated by $*\,p < 0.10$, $**\,p < 0.05$, $***\,p < 0.01$.

result thus appears somewhat puzzling, and bears the question of whether the observed effect on the labor index is in fact driven by individuals' *supply* of labor or is, to some extent, induced by a drop in local labor *demand*.

Loudspeaker announcements were assigned to entire villages and thus may not only have affected wage laborers, but potentially also their local employers. The latter, in turn, might have responded to the awareness campaign by reducing their demand for labor. As our sample is mainly composed of daily wage laborers, and contains only a small number of potential employers, we are unable to directly measure supply and demand effects in our data. However, we conduct several additional analyses to shed light on the underlying mechanisms to the extent possible with the data at hand. First, we look at the effects of the two treatments on the wage rate. If there was indeed a shortage in labor supply, we would—under the assumption that labor is immobile across villages—expect that wage rates increased in the *phone plus loudspeaker* group, but not in the other experimental groups. If, instead, the effect is driven by a lack of demand for labor, we would expect the surplus of laborers to negatively affect wage rates. While we find some indication for an increase in wages among those who received both the awareness call and loudspeaker messages, the effects are not significantly different from zero.

In a second step, we consider heterogeneity along individual level characteristics other than labor types as an indicator for whether we observe a supply or a demand side effect. Intuitively, if the driving force behind the reported labor adjustment is actually a drop in labor demand,

non-work-related individual characteristics should be uncorrelated with the treatment effect. An individual level characteristic of major interest in the context of an effect driven by messages provided through an Imam is religion.

Heterogeneous Effects by Religiosity   As illustrated in the previous section, the effects of our awareness campaign are entirely driven by the loudspeaker announcements made by Imams. Based on our hypothesis that this is due to the level of trust in religious leaders, one would expect that these effects are higher among more, than among less religious individuals. In addition, empirical evidence suggests that religiosity predicts pro-sociality, another characteristic that may correlate with a higher likelihood of adopting behavior that could benefit one's own health, as well as the health of others (Kelly et al., 2024).

We measure religiosity using the number of days on which a mosque or church was visited in the seven days prior to the baseline interview. A man is considered to be *highly religious* when he visited the mosque more often than the average man in our sample, i.e., on at least four days.[26] Note that, while indicating the frequency with which a mosque was visited, higher religiosity is not expected to significantly predict the intensity of treatment. In fact, loudspeaker announcements at mosques are usually heard throughout the entire village and, in case of larger villages, announcements through multiple mosques guarantee larger coverage.

Table 2.8 displays the estimated effects on the labor supply, social interactions, and preventive measures indices for male individuals with high versus low religiosity.[27] In line with our hypothesis, we detect a significant effect for the *phone plus loudspeaker* treatment on the labor supply index for highly religious men only. At the same time, we find no evidence of highly religious male individuals reducing their social interactions. On the contrary, the downward adjustment of social activities in response to the awareness campaign seems to be driven by less religious men. As shown in Column (4), men who reported less than four mosque visits in the seven days prior to the baseline interview reduced social interactions by 0.07 standard deviations in response to the *phone plus loudspeaker* treatment. This suggests that individuals with both high and low levels of religiosity value information delivered by a religious leader, but choose different behavioral responses to decrease the risk of getting infected or infecting others. For instance, their valuation of religious gatherings may make highly religious men on average less willing to reduce social interactions (including mosque visits). Comparing the point estimates across the standardized indices does indicate, however, that the extent to which individuals adjust behavior is indeed positively correlated with the level of religiosity.

---

[26] The average is significantly lower for women, for whom two days of visits are above the average. Taking into account both women and men, high religiosity is defined as having visited the mosque or church on at least three days.

[27] Tables 2.B.11 and 2.B.12 in the Appendix display the corresponding estimates for the overall sample and female individuals, respectively.

Table 2.8: Effects on Labor Supply, Social Interactions, and Preventive Measures by Religiosity: Male Subsample

| | Less Labor Supply | | Less Social Interactions | | More Preventive Measures | |
	High Religiosity (1)	Low Religiosity (2)	High Religiosity (3)	Low Religiosity (4)	High Religiosity (5)	Low Religiosity (6)
Phone	-0.004	0.010	0.018	0.012	-0.024	0.032
	(0.043)	(0.040)	(0.021)	(0.022)	(0.028)	(0.025)
Phone + loudsp.	0.152**	0.031	0.011	0.065*	0.007	0.019
	(0.066)	(0.060)	(0.038)	(0.033)	(0.044)	(0.041)
p-val phone=phone+loudsp.	0.015	0.711	0.851	0.105	0.473	0.752
p-val any treatment	0.579	0.727	0.418	0.300	0.479	0.220
Obs.	1,480	1,825	1,480	1,825	1,480	1,825
Villages	466	562	466	562	466	562
R^2	0.401	0.391	0.417	0.382	0.459	0.470
Control mean	-0.243	-0.177	-0.161	-0.116	0.025	-0.035
SD	0.933	0.941	0.455	0.483	0.622	0.567

Notes: The table shows treatment effects on summary indices of labor supply, social interactions, and preventive measures for the male subsample. Results are reported for individuals with high and low levels of religiosity, respectively. Significance levels are indicated by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Alternative Explanations   Given that the *phone plus loudspeaker* treatment differs from the *phone* treatment along several dimensions—in particular the mode of delivery, the frequency of delivery, and the identity of the messenger—it is natural to ask whether the additional effects of the loudspeaker announcements are indeed driven by the religious component of the treatment. While we are unable to perfectly disentangle the different channels through which the treatment could affect perceptions and behavior, we provide several pieces of suggestive evidence that are in line with our interpretation.

The most prominent concern is that loudspeaker announcements may act as a reminder rather than affecting individuals through a trust channel. To investigate the reminding role of loudspeaker announcements, we look at heterogeneity of the *phone plus loudspeaker* treatment effects depending on the frequency with which announcements were made according to our monitoring data. We divide villages that received loudspeaker announcements into quintiles of announcement frequency and look at the effects of the treatment within each quintile. We find no consistent evidence of a relationship between the frequency of loudspeaker announcements made and our outcomes of interest. This is true in particular for male individuals, who show the most substantial responses to the treatment.[28] We also show that there is heterogeneity in the effects of the *phone plus loudspeaker* treatment depending on individual level religiosity, suggesting once more a link between the identity of the messenger and the effect of the treatment. While these results are insufficient to rule out that a reminder effect may in part explain our findings, they provide additional supportive evidence for our hypothesis that trust in the source of information matters.

Finally, information gathered during our endline survey shows that individuals in the control and both treatment groups are on average equally likely to report that they received a call about COVID-19. There is also no significant difference in the likelihood of reporting that they heard loudspeaker announcements about COVID-19 between the *phone* and the *phone plus loudspeaker* group (see Appendix Table 2.B.2 for more details). However, individuals in the *phone plus loudspeaker* group are significantly more likely to indicate that these announcements were made from a mosque. This not only shows that individuals have—potentially repeatedly— received messages about the virus independently of the treatment assignment within our project. It also indicates that the main difference between information received by the *phone* and *phone plus loudspeaker* groups is in fact the identity of the sender, rather than the mode of distribution. Overall, while neither of the alternative channels can be fully excluded, these results alleviate concerns that our findings may be driven entirely by a reminder effect or a distribution mode effect and underline the importance of receiving information through a trusted source.

---

[28] We show the differential effects of the *phone plus loudspeaker* treatment by frequency quintile on the labor supply index in Appendix Figure 2.B.3. Additional results are available upon request.

2.6   Conclusion

In this paper, we explore the role of religious leaders in shaping beliefs and preventive behavior in the context of health crises. We add to the literature by highlighting the importance of religious figures as supporters of government recommendations, thus crucially contributing to their containment strategies. To this end, we conducted a randomized remote awareness campaign endorsed by local religious leaders during the COVID-19 pandemic in rural Pakistan. We study the impact of the campaign on behavioral changes in labor supply and social interactions, as well as on the take-up of preventive measures.

We find significant effects of the treatment on labor supply and social interactions among male individuals who were exposed to both an awareness call and loudspeaker announcements by an Imam. Whereas this effect does not seem to be driven by changes in knowledge, men in the *phone plus loudspeaker* group significantly increased their concerns about transmitting the virus to others—suggesting a likely channel through which the treatment may have affected behavior. Despite significant increases in knowledge, we observe no effects on behavior among female individuals. This may be partly driven by a decrease in the estimated cost of getting infected. While our design does not allow us to causally disentangle the importance of religious leaders as providers of information from a potential reminder effect through additional messages, we provide suggestive evidence in support of the former interpretation.

Our study of the informational mechanism behind the influence of religion on economic and social outcomes opens interesting avenues for future research. Religious leaders simultaneously act as credible information sources and as trusted support providers in case of adverse shocks, suggesting several channels through which they may affect individual behavior. On the one hand, the effects might be driven by institutional credibility of religion-supporting organizations extended to their leaders as representative agents. On the other hand, the social proximity often developed as a result of repeated interactions between local religious leaders and adherents might play a decisive role in the trust building process and thus explain the reported effects. Disentangling these channels represents an important step forward for advancing our knowledge on the informational role of religious leaders in shaping economic performance and social behaviors.

Bibliography

Abaluck, J., Kwong, L. H., Styczynski, A., Haque, A., Kabir, M. A., Bates-Jefferys, E., Craw-ford, E., Benjamin-Chung, J., Raihan, S., Rahman, S., Benhachmi, S., Bintee, N. Z., Winch, P. J., Hossain, M., Reza, H. M., Jaber, A. A., Momen, S. G., Rahman, A., Banti, F. L., Huq, T. S., Luby, S. P., and Mobarak, A. M. (2022). Impact of community masking on COVID-19: A cluster-randomized trial in Bangladesh. *Science*, 375(6577):eabi9069.

Acemoglu, D., Reed, T., and Robinson, J. A. (2014). Chiefs: Economic development and elite control of civil society in Sierra Leone. *Journal of Political Economy*, 122(2):319–368.

Alatas, V., Chandrasekhar, A. G., Mobius, M., Olken, B. A., and Paladines, C. (2024). Do celebrity endorsements matter? A twitter experiment promoting vaccination in Indonesia. *The Economic Journal*, 134(659):913–933.

Armand, A., Augsburg, B., Kameshwara, K. K., and Bancalari, A. (2022). Social proximity and misinformation: Experimental evidence from a mobile phone-based campaign in India. IFS Working Paper 22/18.

Ashraf, N., Field, E., Voena, A., and Ziparo, R. (2023). Gendered spheres of learning and household decision making over fertility. Stanford King Center on Global Development Working Paper No. wp2043.

Ashraf, N., Jack, B. K., and Kamenica, E. (2013). Information and subsidies: Complements or substitutes? *Journal of Economic Behavior & Organization*, 88:133–139.

Auriol, E., Lassebie, J., Panin, A., Raiber, E., and Seabright, P. (2020). God insures those who pay? Formal insurance and religious offerings in Ghana. *The Quarterly Journal of Economics*, 135(4):1799–1848.

Balán, P., Bergeron, A., Tourek, G., and Weigel, J. L. (2022). Local elites as state capacity: How city chiefs use local information to increase tax compliance in the Democratic Republic of the Congo. *American Economic Review*, 112(3):762–797.

Banerjee, A., Alsan, M., Breza, E., Chandrasekhar, A. G., Chowdhury, A., Duflo, E., Goldsmith-Pinkham, P., and Olken, B. A. (2020). Messages on COVID-19 prevention in india increased symptoms reporting and adherence to preventive behaviors among 25 mil-lion recipients with similar effects on non-recipient members of their communities. NBER Working Paper No. w27496.

Banerjee, A., Chandrasekhar, A. G., Dalpath, S., Duflo, E., Floretta, J., Jackson, M. O., Kan-nan, H., Loza, F. N., Sankar, A., Schrimpf, A., et al. (2021). Selecting the most effective nudge: Evidence from a large-scale experiment on immunization. NBER Working Paper No. w28726.

Banerjee, A., Chandrasekhar, A. G., Duflo, E., and Jackson, M. O. (2019). Using gossips to spread information: Theory and evidence from two randomized controlled trials. *The Review of Economic Studies*, 86(6):2453–2490.

Banerjee, A. and Iyer, L. (2005). History, institutions, and economic performance: The legacy of colonial land tenure systems in India. *American Economic Review*, 95(4):1190–1213.

Bardhan, P. (2002). Decentralization of governance and development. *Journal of Economic Perspectives*, 16(4):185–205.

Bardhan, P. and Mookherjee, D. (2006). Decentralisation and accountability in infrastructure delivery in developing countries. *The Economic Journal*, 116(508):101–127.

Bassi, V. and Rasul, I. (2017). Persuasion: A case study of papal influences on fertility-related beliefs and behavior. *American Economic Journal: Applied Economics*, 9(4):250–302.

Basten, C. and Betz, F. (2013). Beyond work ethic: Religion, individual, and political preferences. *American Economic Journal: Economic Policy*, 5(3):67–91.

Beath, A., Christia, F., and Enikolopov, R. (2017). Direct democracy and resource allocation: Experimental evidence from Afghanistan. *Journal of Development Economics*, 124:199–213.

Becker, S. O. and Woessmann, L. (2009). Was Weber wrong? A human capital theory of protestant economic history. *The Quarterly Journal of Economics*, 124(2):531–596.

Belloc, M., Drago, F., and Galbiati, R. (2016). Earthquakes, religion, and transition to self-government in Italian cities. *The Quarterly Journal of Economics*, 131(4):1875–1926.

Ben-Nun Bloom, P., Arikan, G., and Courtemanche, M. (2015). Religious social identity, religious belief, and anti-immigration sentiment. *American Political Science Review*, 109(2):203–221.

Bentzen, J. S. (2021). In crisis, we pray: Religiosity and the COVID-19 pandemic. *Journal of Economic Behavior & Organization*, 192:541–583.

Bryan, G., Choi, J. J., and Karlan, D. (2020). Randomizing religion: The impact of protestant evangelism on economic outcomes. *The Quarterly Journal of Economics*, 136(1):293–380.

Campante, F. and Yanagizawa-Drott, D. (2015). Does religion affect economic growth and happiness? Evidence from Ramadan. *The Quarterly Journal of Economics*, 130(2):615–658.

Cantoni, D. (2015). The economic effects of the protestant reformation: testing the Weber hypothesis in the German lands. *Journal of the European Economic Association*, 13(4):561–598.

Casey, K., Glennerster, R., Miguel, E., and Voors, M. (2023). Skill versus voice in local development. *Review of Economics and Statistics*, 105(2):311–326.

Christensen, D., Dube, O., Haushofer, J., Siddiqi, B., and Voors, M. (2021). Building resilient health systems: Experimental evidence from Sierra Leone and the 2014 Ebola outbreak. *The Quarterly Journal of Economics*, 136(2):1145–1198.

Clingingsmith, D., Khwaja, A. I., and Kremer, M. (2009). Estimating the impact of the Hajj: Religion and tolerance in Islam's global gathering. *The Quarterly Journal of Economics*, 124(3):1133–1170.

Condra, L. N., Isaqzadeh, M., and Linardi, S. (2019). Clerics and scriptures: Experimentally disentangling the influence of religious authority in Afghanistan. *British Journal of Political Science*, 49(2):401–419.

Dasgupta, A. and Beard, V. A. (2007). Community driven development, collective action and elite capture in Indonesia. *Development and Change*, 38(2):229–249.

Dupas, P., Huillery, E., and Seban, J. (2018). Risk information, risk salience, and adolescent sexual behavior: Experimental evidence from Cameroon. *Journal of Economic Behavior & Organization*, 145:151–175.

Freedman, M. (2020). Vote with your rabbi: The electoral effects of religious institutions in Israel. *Electoral Studies*, 68:102241.

Goldstein, M. and Udry, C. (2008). The profits of power: Land rights and agricultural investment in Ghana. *Journal of Political Economy*, 116(6):981–1022.

Greyling, C., Maulit, J. A., Parry, S., Robinson, D., Smith, S., Street, A., and Vitillo, R. (2016). Lessons from the faith-driven response to the West Africa Ebola epidemic. *The Review of Faith & International Affairs*, 14(3):118–123.

Gruber, J. and Hungerman, D. M. (2008). The church versus the mall: What happens when religion faces increased secular competition? *The Quarterly Journal of Economics*, 123(2):831–862.

Gruber, J. H. (2005). Religious market structure, religious participation, and outcomes: Is religion good for you? *The BE Journal of Economic Analysis & Policy*, 5(1):0000101515153806371454.

Haerpfer, C., Inglehart, R., Moreno, A., Welzel, C., Kizilova, K., Diez-Medrano, J., Lagos, M., Norris, P., Ponarin, E., and Puranen (eds.), B. (2022). World values survey: Round seven - country-pooled datafile version 5.0. JD Systems Institute and WVSA Secretariat, Madrid, Spain, and Vienna, Austria. doi:10.14281/18241.20 [last accessed April 10, 2024].

International Labor Organization (2024). ILO modelled estimates database. Available at https://ilostat.ilo.org/data/ [last accessed April 10, 2024].

Jegede, A. S. (2007). What led to the Nigerian boycott of the polio vaccination campaign? *PLoS Medicine*, 4(3):e73.

Kelly, J. M., Kramer, S. R., and Shariff, A. F. (2024). Religiosity predicts prosociality, especially when measured by self-report: A meta-analysis of almost 60 years of research. *Psychological Bulletin*.

Kling, J. R., Liebman, J. B., and Katz, L. F. (2007). Experimental analysis of neighborhood effects. *Econometrica*, 75(1):83–119.

Labonne, J. and Chase, R. S. (2009). Who is at the wheel when communities drive development? Evidence from the Philippines. *World Development*, 37(1):219–231.

Martinez-Bravo, M. and Stegmann, A. (2022). In vaccines we trust? The effects of the CIA's vaccine ruse on immunization in Pakistan. *Journal of the European Economic Association*, 20(1):150–186.

McClendon, G. and Riedl, R. B. (2015). Religion as a stimulant of political participation: Experimental evidence from Nairobi, Kenya. *The Journal of Politics*, 77(4):1045–1057.

Moreno-Medina, J. (2023). Sinning in the rain: Weather shocks, church attendance, and crime. *Review of Economics and Statistics*, 105(1):54–69.

Pakistan Bureau of Statistics (2017a). 2017 population census data. Available online upon request.

Pakistan Bureau of Statistics (2017b). Pakistan National Census Bureau Report 2017. Available at https://www.pbs.gov.pk/sites/default/files/population/census_reports/ncr_pakistan.pdf [last accessed April 21, 2024].

Reinikka, R. and Svensson, J. (2004). Local capture: Evidence from a central government transfer program in Uganda. *The Quarterly Journal of Economics*, 119(2):679–705.

Sánchez de la Sierra, R. (2020). On the origins of the state: Stationary bandits and taxation in eastern Congo. *Journal of Political Economy*, 128(1):32–74.

Sharma, K. (2017). Clerical persuasion and religious extremism: An experiment among Sunni and Shia Muslims in northern India. Unpublished Manuscript, Columbia University.

Solís Arce, J. S., Warren, S. S., Meriggi, N. F., Scacco, A., McMurry, N., Voors, M., Syunyaev, G., Malik, A. A., Aboutajdine, S., Adeojo, O., et al. (2021). COVID-19 vaccine acceptance and hesitancy in low-and middle-income countries. *Nature Medicine*, 27(8):1385–1394.

Sperber, E. and McClendon, G. (2022). Increasing youth political engagement with efficacy not obligation: Evidence from a workshop-based experiment in Zambia. *Political Behavior*, 44(4):1933–1958.

Stroebel, J. and van Benthem, A. (2012). The power of the church-the role of roman catholic teaching in the transmission of hiv. SSRN Working Paper No. 2018071.

Tanaka, S. and Muzones, M. (2016). Policy brief on female labor force participation in Pakistan. ADB Brief No. 70.

Torgler, B. (2006). The importance of faith: Tax morale and religiosity. *Journal of Economic Behavior & Organization*, 61(1):81–109.

Vyborny, K. (2021). Persuasion and public health: Evidence from an experiment with religious leaders during COVID-19 in Pakistan. SSRN Working Paper No. 3842048.

Appendices

2.A Study Details

Survey Instruments We developed survey instruments for baseline and endline interviews and the awareness campaign. Every survey instrument was piloted and enumerators received a training on each of them separately. During the pilot, suggestions and feedback were collected about the questionnaires, including changing translations, correcting filters, rephrasing questions, and reducing the length of the questionnaires. All interviews were conducted via phone by personnel hired by NRSP. Enumerators used computer-assisted surveying to record responses. All survey instruments used simplified terms and were provided in Urdu and Sindhi, the two local languages spoken in the study area.

Consent In prior projects, every household was explicitly asked to consent to the storage of their phone numbers (by NRSP). For our project, respondents were asked to consent to the first interview and the participation in the research study. In a separate question, respondents were asked for consent to be called again for follow-ups. Information was only collected if participants explicitly confirmed that they were willing to participate in the interview, i.e., when they provided informed consent. During the baseline survey, enumerators were instructed to conduct interviews with any available household member that was above 18 and was registered in the databases of National Rural Support Programme (NRSP). The same household member was then re-interviewed at endline.

Baseline Survey During the baseline interviews, we collected information on topics including basic socio-economic characteristics, health status of the respondent and that of the household members, awareness about COVID-19, perceptions about COVID-19, and the preventive measures applied by the respondent and household members. While we inquired information about COVID-19, we also shared key, targeted messages read out to the respondent if they indicated that their answered revealed misconceptions or a lack of knowledge about the COVID-19 disease. Baseline interviews largely took between 15 and 40 minutes.

Remote Awareness Campaign   The awareness campaign in form of a questionnaire contained mainly two parts: (1) Informative messages about COVID-19, which the enumerator read to the respondent and (2) a short test asking questions about the information that was just shared. The awareness call lasted between 10 and 20 minutes.

Endline Survey   The endline survey covered a subset of modules and questions from the baseline questionnaire. This survey included questions on all outcomes analyzed in the impact evaluation and additionally a module on experiences with the interventions implemented in this study. It took between 10 and 25 minutes to be completed.

Estimation Sample   For the final sample, we exlcude all respondents for whom at least one of the three main outcome indices used in the analysis is not available. While this condition slightly decreases the number of observations, it allows to define a consistent sample across the main regressions, making results more comparable.

Phone Call Message   The corona virus is a new respiratory disease originating from a virus that is highly contagious. The most common symptoms are fever, dry cough and tiredness. Most people who are infected may experience mild illness and recover, but the disease can develop more severely and even be fatal for others. In sever cases a person experiences breathing difficulty, chest pain or pressure, or loss of speech or movement. Please keep in mind that even if you feel healthy and do not develop symptoms you may still have contracted corona virus. Some people infected with corona virus show symptoms, but others do not show any symptoms. For this reason, it is impossible for you to recognize if another person is infected or not. Hence, there may be people feeling healthy who spread the virus and infect other people unknowingly. This makes it more difficult to stop the spread of corona virus than that of other diseases: it can also spread without visible symptoms.

The corona virus is mainly transmitted from person to person through two main modalities:

- Firstly, through small respiratory particles produced from a nearby infected person by coughs or exhales, which enter your mouth or nose.

- Secondly, when touching contaminated objects or other surfaces where droplets from an infected person landed and then touching your mouth, nose or eyes.

As a consequence, transmission is especially likely when people are in close contact with one another.

Because corona virus is a new disease, to date, there is no specific medicine recommended to prevent or treat the new corona virus. Also, please keep in mind that no traditional healing

method can cure coronavirus. This means that the current situation of emergency may protract for many months from now. So, what can we do? Prevention. Three simple measures are highly effective in preventing the spread of corona virus: good hygiene, physical distancing from people, and wearing a face mask. These three are equally important measures, they need to be practiced together to be effective.

Let me tell you a bit about each preventive measure, starting with hygiene. First, wash your hands with water and soap. Normal soap is sufficient to kill the virus, without the soap the virus will stay on your hands. Also, you can use normal water, it is not necessary that water is boiled or purified. You should also refrain from touching your eyes, mouth, or face with your hands to prevent the virus from entering your body. Try to wash your hands as often as possible every day. This means at least 5 times per day. For example after you have been outside your household, after touching an item that could have been touched by other people, before, during and after preparing food, before eating, after coughing and sneezing, before and after taking care of an infected person, after defecating, after handling animals or animals' waste, or when hands are visibly dirty. An easy and cheap solution for the house could be to make use of a 'soapy' water bottle. This can be produced out of any plastic bottle, filled with water mixed with soap. It is sufficient to use a small piece of soap for 1.5 liters. A small hole drilled in the cap allows the bottle to be used efficiently.

Because corona virus may spread when touching contaminated surfaces, it is also recommended to routinely clean surfaces frequently touched by you, like your phone, or touched by other people, with soap and water. It is recommended that your personal items such as a water bottle, cup, spoon, or towel are washed before sharing with other people, especially if they are sick. Remember touching a surface or an object that was touched by a person carrying the corona virus mean that you may also get infected.

Let me now talk about the second preventive measure, which is equally important: physical distancing. One infected person may quickly transmit the virus to many healthy people around them. Remember that even people who look healthy may transmit the virus. The virus can travel from one person to another person. For this reason, you should not only refrain from touching others, e.g., from shaking hands, but also avoid visiting social events or public spaces when crowded, especially if indoors. This means that you might have to adjust how you do certain activities like prayer or other religious and non-religious activities. If you sneeze or cough and have the virus, particles travel in the air up to a few meters around you and you may infect a person that is close enough. If you open both your arms wide, this is about two meters: to stay safe, health experts recommend to keep at least two meters distance from others, anywhere outside your home. The virus may also stay in the air for some time after you sneeze or cough. You should make sure to have enough ventilation when at home or any other indoor place. If any member of your household is sick, then you should assist them while maintaining

two meters distance (about 2 arms length) from them until they recover.

In addition, a third measure to prevent the spread of the virus is wearing a face mask covering your mouth and nose. It is important that you wear the mask whenever you are outside the house, or feeling sick. It is, however, especially recommended that you wear it when you are indoors and can not maintain a safe distance of 2 meters. If you use a face mask, you should not touch it while wearing it. Replace it with a new clean mask when it becomes damp. Do not re-use single use masks. When you do not wear a mask, it is recommendable to cover your mouth and nose with your flexed elbow or a tissue or towel/cloth when coughing or sneezing can be effective even when you are not sick. Also do not spit. Tissues should be disposed immediately after use. Towel, clothes or textile masks you use for these purposes should be boiled after use before hanging them to dry. Remember you need to do all three mentioned measures for highest protection level. The wearing of a mask should be used in addition to personal hygiene and social distancing. Alone, each measure will not be able to fully protect you or others around you from getting the virus.

Now, I will explain to you when you should suspect that you or any household member may have been infected with the corona virus and how to act in this case. You may suspect to be infected with corona virus if:

- you show any of the typical corona virus symptoms like fever, dry cough, or tiredness.

- OR you show any of the severe corona virus symptoms such as breathing difficulty, chest pain or pressure, or loss of speech or movement.

- OR you have been told by a public health official that you may be infected.

- OR you have been in contact with someone known to have corona virus.

- OR you have recently been in contact with people who were sick in or outside your household, or with travelers from other provinces or countries.

Usually, symptoms manifest after 5-6 days, but sometimes incubation of symptoms may take up to 14 days.

In case you suspect that you or any household member may be infected with corona virus but you have not been confirmed yet, it is important that you contact the health facilities, community leader or a trusted individual that can help while you stay at home (quarantine) in the meantime. You may also contact the helpline number 1166 put in place by the government of Pakistan. If you are confirmed to have corona virus, you should quarantine until you recover and follow the instructions of the health facility. There are different measures that you may take. You may either call a hospital or health facility or stay at home, isolate yourself from other people, make sure you were a mask, clean your hands often, make sure that other household members

do not share or touch objects you have used and monitor your symptoms closely. By doing so you protect your household and the community, especially the elderly and vulnerable people. Remember in summary, common symptoms are fatigue, dry cough and fever. Sever symptoms are chest pain or pressure, loss of speech or movement, difficulty breathing or shortness of breath. You can protect yourself by following the three measures explained: social distancing, wearing mask and washing hands. For additional measures you may always make sure that you clean all surfaces, avoid visiting public spaces such as market or mosque when crowded, avoid hand shakes, wear disposable gloves, avoid public transport, and protect older and ill people. It is important that you do not listen to rumors but rely on credible sources of information.

Loudspeaker Announcement Message The corona virus is a new respiratory disease originating from a virus that is highly contagious. The typical symptoms are fever, dry cough and difficulty when breathing. Most people who are infected may experience mild illness and recover, but it can be more severe or deadly for others. It may be possible that a person can get corona virus by touching a surface or object that has the virus on it and then touching their own mouth, nose, or possibly their eyes. It is recommended by health experts that you practice "hand hygiene", by washing hands with soap frequently also recommend routine cleaning of frequently touched surfaces like your phone. Avoid large events and mass gatherings as they can contribute to the spread of corona virus. People in attendance at these events may be sick and can transmit the virus when they come in contact with other healthy people. Health experts recommend to maintain at least 2 meters (6 feet) distance to people outside your household. This is because when someone coughs or sneezes they spray small liquid droplets from their nose or mouth which may contain virus. If you are too close, you can breathe in the droplets, including the corona virus, if the person coughing has the disease. In case you suspect that you or one of your household members is infected of a virus, it is important that you contact the health facilities and stay home in the meantime. You may also contact the helpline number 1166 put in place by the government of Pakistan.

2.B Figures and Tables

Table 2.B.1: Randomization Balance of Selected Individual Level Characteristics

Variable	Control Mean/SE (1)	Phone Mean/SE (2)	Phone + Loudspeaker Mean/SE (3)	T-test difference (1) - (2) (4)	(1) - (3) (5)	(2) - (3) (6)
Panel A. Baseline characteristics						
Female	0.521	0.511	0.555	0.010	-0.034	-0.044
	(0.023)	(0.019)	(0.032)			
Age	37.472	37.946	37.876	-0.473	-0.403	0.070
	(0.260)	(0.215)	(0.370)			
Illiterate / below primary education	0.512	0.513	0.489	-0.001	0.023	0.024
	(0.016)	(0.014)	(0.025)			
Wage for work performed outside home (last 7d)	749.598	838.094	738.353	-88.495	11.245	99.741
	(44.518)	(41.097)	(64.533)			
Household size	8.354	8.321	8.140	0.033	0.214	0.181
	(0.129)	(0.096)	(0.165)			
HH owns either land or livestock	0.637	0.620	0.614	0.017	0.023	0.006
	(0.016)	(0.012)	(0.024)			
Number of HH members above 60	0.453	0.458	0.448	-0.005	0.005	0.010
	(0.017)	(0.013)	(0.023)			
Number of HH members severely ill at the beginning of the year	0.204	0.229	0.223	-0.025	-0.019	0.006
	(0.014)	(0.012)	(0.022)			
Number of HH members sick with common COVID-19 symptoms (last 14d)	0.095	0.096	0.076	-0.001	0.019	0.020
	(0.014)	(0.008)	(0.011)			
Panel B. Variables of interest						
Worked outside home (last 7d)	0.351	0.383	0.344	-0.032	0.007	0.039
	(0.018)	(0.015)	(0.026)			
Number of hours worked outside home (yesterday)	2.385	2.468	2.279	-0.083	0.106	0.189
	(0.144)	(0.123)	(0.190)			
Traveled outside village for leisure (last 7d)	0.079	0.073	0.073	0.005	0.006	0.001
	(0.006)	(0.005)	(0.008)			
Had visits for a.l.1 day (last 7d)	0.256	0.271	0.283	-0.016	-0.027	-0.011
	(0.015)	(0.013)	(0.022)			
Practices all three main preventive measures	0.240	0.250	0.260	-0.010	-0.020	-0.009
	(0.017)	(0.014)	(0.024)			
Number of occasions after which respondent washes hands	3.037	3.098	3.174	-0.061	-0.137	-0.076
	(0.086)	(0.082)	(0.136)			
Able to list all three typical symptoms	0.137	0.144	0.146	-0.007	-0.009	-0.002
	(0.015)	(0.013)	(0.023)			
Able to list all three main preventive measures	0.292	0.304	0.316	-0.011	-0.023	-0.012
	(0.018)	(0.015)	(0.027)			
Heard of people treated badly because of their COVID-19 infection	0.079	0.069	0.050	0.010	0.029**	0.019
	(0.011)	(0.008)	(0.009)			
Knows that traditional healers cannot heal a COVID-19 infection	0.443	0.393	0.392	0.050**	0.051	0.001
	(0.019)	(0.016)	(0.028)			
Would like to get a COVID-19 test	0.590	0.584	0.560	0.006	0.030	0.024
	(0.016)	(0.013)	(0.023)			
Has moderate or high concerns of getting infected	0.313	0.309	0.269	0.004	0.044	0.040
	(0.021)	(0.018)	(0.027)			
Has moderate or high concerns of infecting others	0.437	0.423	0.405	0.014	0.032	0.018
	(0.024)	(0.020)	(0.033)			
Perceived average cost of getting infected	12577.717	17752.138	14365.035	-5174.422**	-1787.318	3387.104
	(1457.767)	(1678.042)	(2523.996)			
Thinks that COVID-19 is a severe disease	0.489	0.509	0.519	-0.020	-0.029	-0.010
	(0.023)	(0.020)	(0.033)			
All or some HH members have access to masks	0.745	0.746	0.780	-0.001	-0.035	-0.034
	(0.016)	(0.013)	(0.019)			
Received info on COVID-19 via NGO or mosque loudspeaker	0.108	0.103	0.119	0.005	-0.011	-0.016
	(0.011)	(0.009)	(0.016)			
Panel C. Data collection characteristics						
Individual part of first sample	0.797	0.807	0.843	-0.010	-0.047	-0.037
	(0.025)	(0.020)	(0.030)			
Individual with more than 5 imputations at baseline	0.110	0.100	0.111	0.010	-0.001	-0.011
	(0.009)	(0.007)	(0.013)			

Notes: The table shows randomization balance for selected individual level characteristics at baseline. Columns (1), (2), and (3) report means and standard errors among individuals in villages assigned to the control group, villages assigned to phone calls, and villages assigned to phone calls and loudspeaker announcements, respectively. Columns (4), (5), and (6) report t-tests for differences in means. Significance levels are indicated by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 2.B.2: Self-reported Take-up and Content of Awareness Calls and Loudspeaker Messages

	Control		Phone		Phone + Loudspeaker		T-test difference		
Variable	N/Clusters	Mean/SE	N/Clusters	Mean/SE	N/Clusters	Mean/SE	(1) - (2)	(1) - (3)	(2) - (3)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A: Awareness calls									
Individual received information call about COVID-19	2153 (295)	0.860 (0.008)	3136 (445)	0.871 (0.006)	1092 (148)	0.860 (0.012)	-0.011	0.000	0.011
...on the symptoms, transmission, and preventive measures	1838 (293)	0.950 (0.005)	2710 (443)	0.965 (0.003)	933 (146)	0.973 (0.006)	-0.014**	-0.023***	-0.009
...on the severity of the virus	1828 (293)	0.907 (0.007)	2693 (443)	0.905 (0.005)	931 (146)	0.918 (0.010)	0.002	-0.011	-0.013
...on the cost of contracting the virus	1803 (293)	0.783 (0.009)	2667 (443)	0.792 (0.008)	920 (146)	0.796 (0.013)	-0.009	-0.013	-0.004
...on the likelihood of contracting the virus	1832 (293)	0.930 (0.006)	2699 (443)	0.934 (0.005)	933 (146)	0.950 (0.008)	-0.004	-0.019*	-0.016*
...on the likelihood of infecting others	1829 (293)	0.923 (0.006)	2693 (443)	0.931 (0.005)	924 (146)	0.934 (0.008)	-0.007	-0.011	-0.003
...asking for the health status of HH members	1838 (293)	0.970 (0.004)	2713 (443)	0.976 (0.003)	933 (146)	0.977 (0.006)	-0.006	-0.008	-0.002
...asking for help in spreading info about the virus in one's community	1813 (293)	0.881 (0.009)	2675 (443)	0.898 (0.006)	921 (146)	0.907 (0.010)	-0.017	-0.025*	-0.009
Shared content of the call with others	1850 (293)	0.925 (0.006)	2726 (443)	0.925 (0.005)	938 (146)	0.938 (0.008)	0.001	-0.013	-0.013
HH member received information call about COVID-19	1727 (293)	0.442 (0.013)	2557 (442)	0.437 (0.010)	867 (145)	0.419 (0.017)	0.005	0.024	0.019
Panel B: Loudspeaker messages									
Heard loudspeaker announcement informing about COVID-19	1753 (293)	0.441 (0.012)	2596 (443)	0.456 (0.010)	897 (145)	0.482 (0.018)	-0.015	-0.041*	-0.026
Announcement was made at a mosque	769 (261)	0.843 (0.013)	1174 (395)	0.834 (0.011)	431 (132)	0.884 (0.016)	0.009	-0.041**	-0.050***

Notes: The table shows self-reported variables capturing the take-up awareness phone calls (Panel A) and the loudspeaker announcements (Panel B). Columns (1), (3), and (5) report the number of observations and clusters in villages assigned to the control group, villages assigned to phone calls, and villages assigned to phone calls and loudspeaker announcements, respectively. Columns (2), (4), and (6) show the respective mean values and standard errors. Columns (7), (8), and (9) report t-tests for differences in means. Significance levels are indicated by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 2.B.3: Balance of Selected Individual Level Characteristics—Final Sample

Variable	Control Mean/SE (1)	Phone Mean/SE (2)	Phone + Loudspeaker Mean/SE (3)	T-test difference (1) - (2) (4)	(1) - (3) (5)	(2) - (3) (6)
Panel A. Baseline characteristics						
Female	0.478	0.476	0.522	0.002	-0.044	-0.046
	(0.025)	(0.020)	(0.035)			
Age	37.878	38.648	38.087	-0.770*	-0.209	0.561
	(0.306)	(0.253)	(0.432)			
Illiterate / below primary education	0.503	0.508	0.473	-0.005	0.030	0.035
	(0.017)	(0.014)	(0.027)			
Wage for work performed outside home (last 7d)	778.255	871.414	793.327	-93.159	-15.071	78.087
	(48.296)	(44.958)	(71.070)			
Household size	8.458	8.451	8.263	0.007	0.196	0.189
	(0.134)	(0.107)	(0.192)			
HH owns either land or livestock	0.651	0.634	0.624	0.017	0.027	0.010
	(0.017)	(0.013)	(0.026)			
Number of HH members above 60	0.446	0.445	0.432	0.001	0.014	0.013
	(0.019)	(0.015)	(0.025)			
Number of HH members severely ill at the beginning of the year	0.212	0.236	0.247	-0.023	-0.034	-0.011
	(0.015)	(0.014)	(0.026)			
Number of HH members sick with common COVID-19 symptoms (last 14d)	0.089	0.101	0.088	-0.012	0.001	0.013
	(0.011)	(0.010)	(0.013)			
Panel B. Variables of interest						
Worked outside home (last 7d)	0.363	0.397	0.364	-0.035	-0.001	0.033
	(0.019)	(0.017)	(0.029)			
Number of hours worked outside home (yesterday)	2.519	2.593	2.496	-0.074	0.023	0.097
	(0.162)	(0.134)	(0.228)			
Traveled outside village for leisure (last 7d)	0.074	0.079	0.078	-0.004	-0.004	0.001
	(0.008)	(0.006)	(0.010)			
Had visits for a.l.1 day (last 7d)	0.249	0.273	0.278	-0.024	-0.029	-0.005
	(0.016)	(0.014)	(0.025)			
Practices all three main preventive measures	0.262	0.264	0.283	-0.002	-0.021	-0.019
	(0.019)	(0.016)	(0.029)			
Number of occasions after which respondent washes hands	3.147	3.200	3.306	-0.053	-0.159	-0.107
	(0.095)	(0.091)	(0.145)			
Able to list all three typical symptoms	0.146	0.143	0.144	0.003	0.002	-0.001
	(0.017)	(0.013)	(0.025)			
Able to list all three main preventive measures	0.316	0.322	0.337	-0.005	-0.021	-0.015
	(0.021)	(0.017)	(0.031)			
Heard of people treated badly because of their COVID-19 infection	0.086	0.077	0.058	0.009	0.027	0.018
	(0.014)	(0.009)	(0.012)			
Knows that traditional healers cannot heal a COVID-19 infection	0.455	0.406	0.400	0.049*	0.055	0.006
	(0.020)	(0.017)	(0.031)			
Would like to get a COVID-19 test	0.600	0.588	0.579	0.012	0.021	0.009
	(0.017)	(0.014)	(0.025)			
Has moderate or high concerns of getting infected	0.292	0.290	0.256	0.001	0.036	0.034
	(0.022)	(0.019)	(0.029)			
Has moderate or high concerns of infecting others	0.411	0.401	0.405	0.010	0.006	-0.004
	(0.025)	(0.021)	(0.035)			
Perceived average cost of getting infected	13825.755	18224.833	15755.040	-4399.078*	-1929.285	2469.794
	(1830.595)	(1839.834)	(3169.798)			
Thinks that COVID-19 is a severe disease	0.493	0.513	0.517	-0.020	-0.024	-0.004
	(0.026)	(0.021)	(0.036)			
All or some HH members have access to masks	0.773	0.757	0.783	0.016	-0.009	-0.025
	(0.016)	(0.014)	(0.022)			
Received info on COVID-19 via NGO or mosque loudspeaker	0.100	0.096	0.101	0.003	-0.001	-0.004
	(0.011)	(0.009)	(0.016)			
Panel C. Data collection characteristics						
Individual part of first sample	0.791	0.783	0.813	0.008	-0.022	-0.030
	(0.027)	(0.023)	(0.038)			
Individual with more than 5 imputations at baseline	0.064	0.052	0.058	0.011	0.006	-0.005
	(0.010)	(0.007)	(0.013)			

Notes: The table shows balance for selected individual level characteristics at baseline for the final analysis sample. Columns (1), (2), and (3) report means and standard errors among individuals in villages assigned to the control group, villages assigned to phone calls, and villages assigned to phone calls and loudspeaker announcements, respectively. Columns (4), (5), and (6) report t-tests for differences in means. Significance levels are indicated by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 2.B.4: Baseline Characteristics by Gender: Male and Female Subsamples

	Mean (1)	SD (2)	Min. (3)	Max. (4)	# (5)
Panel B: Awareness					
Male					
Knows symptoms of corona virus	0.88	0.33	0	1	3,219
Frequently and thoroughly wash hands with soap	0.82	0.38	0	1	3,219
Wear a face mask	0.61	0.49	0	1	3,219
Maintain two meters distance from people	0.53	0.50	0	1	3,219
Female					
Knows symptoms of corona virus	0.84	0.36	0	1	2,898
Frequently and thoroughly wash hands with soap	0.77	0.42	0	1	2,900
Wear a face mask	0.54	0.50	0	1	2,900
Maintain two meters distance from people	0.47	0.50	0	1	2,900
Panel C: Perceptions					
Male					
Thinks one would die or recover with severe health damages if infected	0.51	0.50	0	1	2,319
Has moderate or high concerns of getting infected	0.20	0.40	0	1	3,147
Has moderate or high concerns of infecting other household members	0.33	0.47	0	1	3,111
Female					
Thinks one would die or recover with severe health damages if infected	0.51	0.50	0	1	2,352
Has moderate or high concerns of getting infected	0.39	0.49	0	1	2,851
Has moderate or high concerns of infecting other household members	0.49	0.50	0	1	2,772
Panel D: Behavior					
Male					
Frequently and thoroughly washes hands with soap	0.79	0.41	0	1	3,033
Not touching any objects used by other people before washing them with soap	0.14	0.34	0	1	3,037
Maintains two meters distance from people	0.43	0.49	0	1	3,036
Wears a face mask	0.50	0.50	0	1	2,895
Female					
Frequently and thoroughly washes hands with soap	0.73	0.44	0	1	2,612
Not touching any objects used by other people before washing them with soap	0.11	0.31	0	1	2,622
Maintains two meters distance from people	0.37	0.48	0	1	2,617
Wears a face mask	0.40	0.49	0	1	2,410
Male					
Left village (last 7 days)	0.23	0.42	0	1	3,364
Went to the market (last 7 days) (resp. or household member)	0.70	0.46	0	1	3,345
Went to the mosque, church or mandir (last 7 days)	0.75	0.43	0	1	3,305
Female					
Left village (last 7 days)	0.23	0.42	0	1	3,112
Went to the market (last 7 days) (resp. or household member)	0.72	0.45	0	1	3,101
Went to the mosque, church or mandir (last 7 days)	0.35	0.48	0	1	2,892
Male					
Worked outside home (last 7 days)	0.50	0.50	0	1	3,370
# of days worked outside home (last 7 days, cond.)	5.11	2.19	0	7	1,629
# of hours worked outside home yesterday (cond.)	6.96	3.46	0	14	1,646
Female					
Worked outside home (last 7 days)	0.26	0.44	0	1	3,113
# of days worked outside home (last 7 days, cond.)	4.69	2.39	0	7	759
# of hours worked outside home yesterday (cond.)	6.37	3.79	0	14	778

Notes: The table reports summary statistics for the male and female subsample. Missing observations are due to respondents indicating that they do not know the answer to a question or do not wish to reply to a question (overall sample sizes: N men = 3,370; N women = 3,113).

Table 2.B.5: Component Variables by Index

Index	Indicators
Labor supply	- Has worked outside home in last 7 days
	- Number of days worked outside home in last 7 days
	- Number of hours worked outside home yesterday
Social interactions	- Has received visits for at least one day in last 7 days
	- Number of days with visits in last 7 days
	- Went to the mosque, church or mandir in last 7 days
	- Number of days went to the mosque, church, or mandir in last 7 days
	- Attended any social gathering in last 7 days
	- Had at least one social contact outside home in last 7 days
	- Number of social interactions not restricted in last 7 days
	- Did not restrict social contact for at least 1 day
	- Number of days not restricted social contact
	- Received visits for at least 1 day in the past 7 days from another town
	- Respondent or HH member went to the market in last 7 days
Preventive measures[1]	- Practices the three common preventive measures
	- Number of preventive measures practiced
	- At least some HH member wore mask when leaving home in the last 7 days
	- Number of occasions after which individual washes hands
Knowledge[2]	- Reports to know the symptoms of the COVID-19 virus
	- Correctly reports the three most common symptoms of COVID-19 (fever, dry cough, fatigue)
	- Number of correctly reported symptoms
	- Correctly reports the three most common preventive measures (hand-washing, mask-wearing, distancing)
	- Number of correctly reported preventive measures
	- Knows that an asymptomatic person can spread the virus
	- Knows that one can contract the virus by touching contaminated surfaces
	- Number of correctly indicated transmission channels (out of the above two)
	- Number of correctly reported emergency measures when infection is suspected
Perceived infection risk	- Has moderate or high concerns of getting infected
Perceived transmission risk	- Has moderate or high concerns of infecting other household members
Perceived cost	- Perceived total cost of getting infected
	- Perceived cost of being ill for one week
	- Perceived number of weeks required for recovery if infected
	- Perceived cost of one week treatment or medication if infected
Perceived severity	- Thinks one would die or recover with severe health damages if infected

Notes: The table reports the components of each outcome index.

[1] Other than pre-specified, we do not include an indicator for whether individuals have changed their way of participating in religious congregational activities due to COVID-19 (e.g., by distancing themselves from others). This is due to the fact that the relevant information to define this variable was not collected at endline.

[2] Other than pre-specified, the knowledge index does not contain an indicator for whether an individual has previously heard about the COVID-19 virus. This is due to the nearly universal awareness of COVID-19 at endline and its insufficient value in characterizing knowledge about the virus.

Table 2.B.6: Changes in the Use of Preventive Measures across Treatment Arms (BL to EL)

Variable	Control Villages		Phone Call (only) Villages		Phone Call + Loudsp. Villages		T-test (1)-(2)	T-test (1)-(3)	T-test (2)-(3)
	N/Clusters (1)	Mean/(SE)	N/Clusters (2)	Mean/(SE)	N/Clusters (3)	Mean/(SE)	Mean difference (4)	Mean difference (5)	Mean difference (6)
Panel A. Overall sample									
Δ Hand washing	1627 [287]	0.077 [0.020]	2354 [421]	0.049 [0.017]	815 [141]	0.065 [0.024]	0.028	0.012	-0.016
Δ Distance	1633 [287]	0.072 [0.025]	2355 [421]	0.079 [0.021]	818 [141]	0.100 [0.038]	-0.007	-0.028	-0.021
Δ Mask wearing	1485 [287]	0.207 [0.027]	2151 [418]	0.190 [0.021]	754 [139]	0.196 [0.036]	0.017	0.010	-0.007
Δ Practices the 3 common preventive measures	1897 [292]	0.059 [0.022]	2816 [441]	0.068 [0.018]	980 [146]	0.059 [0.032]	-0.009	-0.000	0.009
Δ # of preventive measures practiced	1896 [292]	0.580 [0.096]	2815 [441]	0.557 [0.085]	980 [146]	0.628 [0.145]	0.023	-0.048	-0.071
Δ A.l. 1 HHm wore mask when leaving home (last 7d)	1569 [277]	0.061 [0.011]	2234 [411]	0.073 [0.011]	819 [141]	0.056 [0.013]	-0.012	0.005	0.017
Panel B. Male subssample									
Δ Hand washing	858 [205]	0.047 [0.025]	1257 [307]	0.036 [0.023]	408 [97]	0.037 [0.027]	0.011	0.010	-0.001
Δ Distance	859 [205]	0.050 [0.035]	1259 [307]	0.061 [0.027]	410 [97]	0.107 [0.048]	-0.011	-0.057	-0.046
Δ Mask wearing	806 [203]	0.186 [0.036]	1170 [299]	0.154 [0.029]	380 [96]	0.145 [0.041]	0.032	0.041	0.009
Δ Practices the 3 common preventive measures	984 [213]	0.028 [0.032]	1479 [327]	0.042 [0.026]	489 [102]	0.047 [0.043]	-0.013	-0.019	-0.005
Δ # of preventive measures practiced	984 [213]	0.397 [0.126]	1479 [327]	0.401 [0.109]	489 [102]	0.507 [0.176]	-0.004	-0.110	-0.106
Δ A.l. 1 HHm wore mask when leaving home (last 7d)	909 [207]	0.059 [0.013]	1273 [304]	0.082 [0.014]	430 [97]	0.058 [0.017]	-0.022	0.001	0.024
Panel C. Female subssample									
Δ Hand washing	769 [224]	0.111 [0.029]	1097 [323]	0.064 [0.023]	407 [111]	0.093 [0.039]	0.047	0.017	-0.030
Δ Distance	774 [224]	0.097 [0.034]	1096 [323]	0.100 [0.029]	408 [111]	0.093 [0.054]	-0.003	0.004	0.007
Δ Mask wearing	679 [220]	0.231 [0.036]	981 [317]	0.232 [0.027]	374 [106]	0.249 [0.052]	-0.001	-0.017	-0.016
Δ Practices the 3 common preventive measures	913 [239]	0.092 [0.026]	1337 [348]	0.097 [0.024]	491 [119]	0.071 [0.042]	-0.005	0.021	0.026
Δ # of preventive measures practiced	912 [239]	0.776 [0.122]	1336 [348]	0.729 [0.121]	491 [119]	0.747 [0.207]	0.047	0.029	-0.018
Δ A.l. 1 HHm wore mask when leaving home (last 7d)	660 [211]	0.064 [0.015]	961 [307]	0.061 [0.013]	389 [112]	0.054 [0.017]	0.002	0.010	0.007

Notes: The table shows changes in the use preventive measures between baseline and endline across treatment arms for the overall sample (Panel A), the male subsample (Panel B), and the female subsample (Panel C). Columns (1), (2), and (3) report the number of observations and clusters, and the mean and standard deviation in villages assigned to the control group, villages assigned to phone calls, and villages assigned to phone calls and loudspeaker announcements, respectively. Columns (4), (5), and (6) report the t-test for differences in mean values. Significance levels are indicated by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 2.B.7: Robustness—Overall sample

	Main specification (1)	Posters (2)	(2) + unbalanced (3)	(3) + T message (4)	Main spec. + T message (5)
Panel A. Behavior					
Less Labor Supply					
Phone	0.008	0.013	0.008	0.017	0.013
	(0.022)	(0.024)	(0.024)	(0.033)	(0.031)
Phone + loudsp.	0.060**	0.078**	0.079**	0.087**	0.064*
	(0.031)	(0.032)	(0.032)	(0.040)	(0.038)
Less Social Interactions					
Phone	0.008	-0.000	0.002	0.014	0.019
	(0.012)	(0.014)	(0.014)	(0.020)	(0.018)
Phone + loudsp.	0.014	0.023	0.024	0.036	0.025
	(0.017)	(0.019)	(0.020)	(0.024)	(0.022)
More Preventive Measures					
Phone	0.014	0.012	0.013	0.005	0.007
	(0.014)	(0.014)	(0.014)	(0.020)	(0.020)
Phone + loudsp.	0.014	0.020	0.021	0.013	0.007
	(0.024)	(0.025)	(0.025)	(0.027)	(0.026)
Panel B. Knowledge					
More Knowledge					
Phone	0.020	0.018	0.019	0.011	0.013
	(0.013)	(0.013)	(0.013)	(0.019)	(0.019)
Phone + loudsp.	0.042**	0.041**	0.042**	0.034	0.034
	(0.019)	(0.021)	(0.021)	(0.024)	(0.023)
Panel C. Beliefs					
Higher Perceived Risk of Getting Infected					
Phone	0.010	0.011	0.011	0.021	0.020
	(0.010)	(0.010)	(0.010)	(0.015)	(0.015)
Phone + loudsp.	0.011	0.013	0.013	0.023	0.022
	(0.016)	(0.016)	(0.016)	(0.018)	(0.018)
Higher Perceived Risk of Transmission					
Phone	0.021**	0.019**	0.020**	0.030**	0.032**
	(0.009)	(0.009)	(0.009)	(0.013)	(0.013)
Phone + loudsp.	0.032**	0.033**	0.033**	0.043**	0.044**
	(0.015)	(0.014)	(0.014)	(0.018)	(0.019)
Higher Perceived Cost					
Phone	0.025	0.021	0.021	-0.011	-0.009
	(0.029)	(0.028)	(0.028)	(0.034)	(0.035)
Phone + loudsp.	-0.044	-0.038	-0.039	-0.072	-0.079*
	(0.038)	(0.039)	(0.039)	(0.045)	(0.045)
Higher Perceived Severity					
Phone	-0.009	-0.009	-0.009	-0.017	-0.017
	(0.013)	(0.013)	(0.013)	(0.019)	(0.019)
Phone + loudsp.	-0.018	-0.016	-0.015	-0.023	-0.026
	(0.019)	(0.019)	(0.019)	(0.023)	(0.024)

Notes: The table shows treatment effects on the main outcomes of interest using several different specifications. Column (1) shows the base specification presented throughout the main paper, including individual and village level characteristics, the baseline level of the outcome of interest, enumerator fixed effects, and an indicator for the poster treatment. Column (2) controls for the outcome at baseline and the poster treatment, and employs enumerator fixed effects. Column (3) augments the specification in Column (2) by adding individual level characteristics that are unbalanced across the three experimental groups. Column (4) adds a control for the individual level assignment to a specific awareness message during the phone call. Finally, Column (5) uses all controls from the main specification and adds to this the indicator for the specific treatment message. Significance levels are indicated by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 2.B.8: Robustness—Male subsample

|	Main specification (1)	Posters (2)	(2) + unbalanced (3)	(3) + T message (4)	Main spec. + T message (5)
Panel A. Behavior					
Less Labor Supply					
Phone	0.010	0.022	0.018	-0.029	-0.034
	(0.028)	(0.029)	(0.029)	(0.044)	(0.044)
Phone + loudsp.	0.089**	0.086**	0.087**	0.044	0.048
	(0.044)	(0.043)	(0.042)	(0.054)	(0.057)
Less Social Interactions					
Phone	0.007	0.005	0.006	0.008	0.010
	(0.015)	(0.016)	(0.015)	(0.024)	(0.024)
Phone + loudsp.	0.044*	0.046*	0.044*	0.046	0.048
	(0.024)	(0.026)	(0.026)	(0.030)	(0.029)
More Preventive Measures					
Phone	0.003	-0.000	0.001	-0.029	-0.025
	(0.017)	(0.017)	(0.017)	(0.025)	(0.025)
Phone + loudsp.	0.011	0.008	0.009	-0.019	-0.016
	(0.030)	(0.030)	(0.030)	(0.034)	(0.034)
Panel B. Knowledge					
More Knowledge					
Phone	0.003	-0.001	-0.001	-0.022	-0.018
	(0.019)	(0.019)	(0.019)	(0.029)	(0.029)
Phone + loudsp.	0.042	0.038	0.038	0.018	0.022
	(0.027)	(0.026)	(0.026)	(0.033)	(0.034)
Panel C. Beliefs					
Higher Perceived Risk of Getting Infected					
Phone	0.012	0.013	0.013	0.014	0.012
	(0.013)	(0.013)	(0.013)	(0.019)	(0.019)
Phone + loudsp.	0.022	0.019	0.020	0.021	0.022
	(0.021)	(0.021)	(0.022)	(0.026)	(0.025)
Higher Perceived Risk of Transmission					
Phone	0.015	0.016	0.015	0.014	0.016
	(0.011)	(0.010)	(0.010)	(0.017)	(0.017)
Phone + loudsp.	0.043**	0.041**	0.041**	0.040	0.045*
	(0.020)	(0.020)	(0.020)	(0.024)	(0.025)
Higher Perceived Cost					
Phone	0.010	0.009	0.006	-0.031	-0.032
	(0.041)	(0.040)	(0.040)	(0.051)	(0.054)
Phone + loudsp.	0.020	0.024	0.028	-0.012	-0.024
	(0.066)	(0.068)	(0.067)	(0.073)	(0.072)
Higher Perceived Severity					
Phone	-0.014	-0.011	-0.011	-0.031	-0.030
	(0.020)	(0.018)	(0.018)	(0.028)	(0.030)
Phone + loudsp.	-0.006	-0.015	-0.015	-0.035	-0.021
	(0.026)	(0.026)	(0.026)	(0.032)	(0.033)

Notes: The table shows treatment effects on the main outcomes of interest using several different specifications. Column (1) shows the base specification presented throughout the main paper, including individual and village level characteristics, the baseline level of the outcome of interest, enumerator fixed effects, and an indicator for the poster treatment. Column (2) controls for the outcome at baseline and the poster treatment, and employs enumerator fixed effects. Column (3) augments the specification in Column (2) by adding individual level characteristics that are unbalanced across the three experimental groups. Column (4) adds a control for the individual level assignment to a specific awareness message during the phone call. Finally, Column (5) uses all controls from the main specification and adds to this the indicator for the specific treatment message. Significance levels are indicated by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 2.B.9: Robustness—Female subsample

	Main specification (1)	Posters (2)	(2) + unbalanced (3)	(3) + T message (4)	Main spec. + T message (5)
Panel A. Behavior					
Less Labor Supply					
Phone	-0.011	0.008	0.002	0.057	0.035
	(0.030)	(0.030)	(0.030)	(0.041)	(0.041)
Phone + loudsp.	0.049	0.060	0.060	0.115**	0.096*
	(0.044)	(0.043)	(0.043)	(0.052)	(0.052)
Less Social Interactions					
Phone	-0.000	0.008	0.010	0.023	0.014
	(0.018)	(0.018)	(0.019)	(0.027)	(0.027)
Phone + loudsp.	-0.007	-0.001	0.001	0.015	0.008
	(0.025)	(0.024)	(0.024)	(0.031)	(0.032)
More Preventive Measures					
Phone	0.016	0.024	0.026	0.045	0.031
	(0.022)	(0.022)	(0.022)	(0.030)	(0.030)
Phone + loudsp.	0.010	0.034	0.037	0.055	0.024
	(0.035)	(0.040)	(0.040)	(0.043)	(0.039)
Panel B. Knowledge					
More Knowledge					
Phone	0.033*	0.028	0.031*	0.035	0.035
	(0.018)	(0.019)	(0.019)	(0.027)	(0.026)
Phone + loudsp.	0.053*	0.047	0.050	0.054	0.055*
	(0.027)	(0.031)	(0.031)	(0.035)	(0.033)
Panel C. Beliefs					
Higher Perceived Risk of Getting Infected					
Phone	0.001	0.007	0.008	0.032	0.021
	(0.016)	(0.016)	(0.016)	(0.022)	(0.023)
Phone + loudsp.	-0.007	0.009	0.009	0.032	0.014
	(0.024)	(0.024)	(0.024)	(0.027)	(0.028)
Higher Perceived Risk of Transmission					
Phone	0.017	0.023	0.024	0.047**	0.039*
	(0.015)	(0.015)	(0.015)	(0.020)	(0.021)
Phone + loudsp.	0.010	0.023	0.023	0.047*	0.033
	(0.022)	(0.020)	(0.020)	(0.025)	(0.027)
Higher Perceived Cost					
Phone	0.032	0.025	0.031	0.001	0.002
	(0.044)	(0.043)	(0.043)	(0.048)	(0.049)
Phone + loudsp.	-0.114***	-0.110***	-0.105***	-0.135**	-0.141**
	(0.042)	(0.038)	(0.039)	(0.055)	(0.056)
Higher Perceived Severity					
Phone	-0.011	-0.002	-0.004	0.000	-0.015
	(0.019)	(0.019)	(0.019)	(0.027)	(0.028)
Phone + loudsp.	-0.016	-0.004	-0.002	0.002	-0.020
	(0.030)	(0.030)	(0.030)	(0.037)	(0.037)

Notes: The table shows treatment effects on the main outcomes of interest using several different specifications. Column (1) shows the base specification presented throughout the main paper, including individual and village level characteristics, the baseline level of the outcome of interest, enumerator fixed effects, and an indicator for the poster treatment. Column (2) controls for the outcome at baseline and the poster treatment, and employs enumerator fixed effects. Column (3) augments the specification in Column (2) by adding individual level characteristics that are unbalanced across the three experimental groups. Column (4) adds a control for the individual level assignment to a specific awareness message during the phone call. Finally, Column (5) uses all controls from the main specification and adds to this the indicator for the specific treatment message. Significance levels are indicated by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Figure 2.B.1: Effects on Single Outcomes of Labor Supply Index: Female Subsample
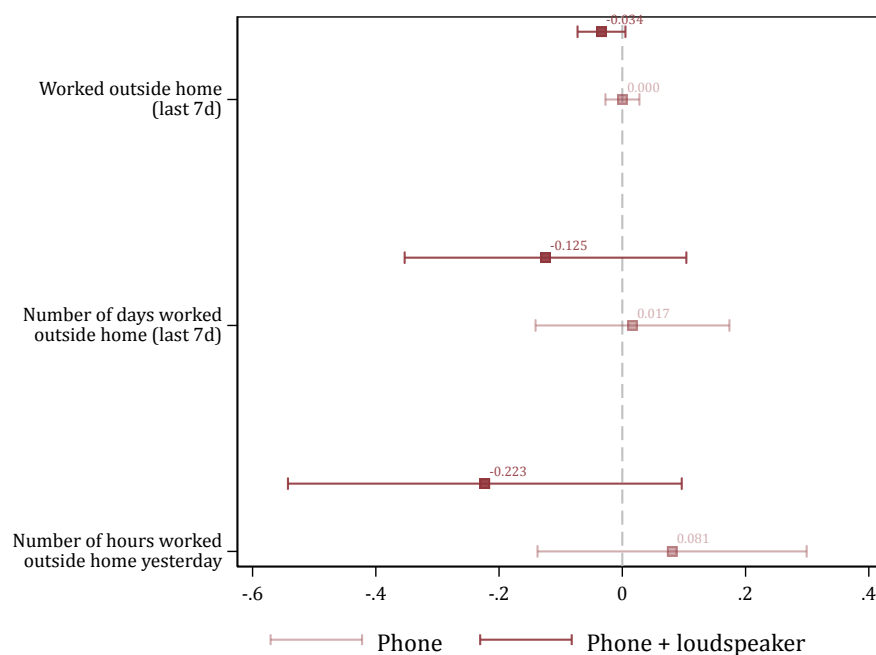
Notes: The graph shows treatment effects on the components of the labor supply index for the female subsample. Point estimates are shown with 10% confidence intervals. Significance levels are indicated by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 2.B.10: Effects on Labor Supply by Employment Status: Overall Sample and Female Subsample

| | | | Less Labor Supply | | | |
	No Work at BL		Daily Wage Laborer		Other Employment	
	All	Female	All	Female	All	Female
	(1)	(2)	(3)	(4)	(5)	(6)
Phone	0.023	-0.005	-0.048	-0.070	-0.020	0.067
	(0.029)	(0.036)	(0.043)	(0.074)	(0.059)	(0.132)
Phone + loudsp.	0.059	0.056	0.155**	0.022	0.002	0.215
	(0.037)	(0.054)	(0.067)	(0.093)	(0.093)	(0.206)
p-val phone=phone+loudsp.	0.303	0.216	0.002	0.317	0.801	0.428
p-val any treatment	0.282	0.857	0.747	0.477	0.778	0.495
Obs.	3,734	2,125	1,664	654	1,085	334
Villages	815	629	555	295	454	217
R^2	0.364	0.421	0.382	0.466	0.472	0.554
Control mean	0.068	0.234	-0.103	0.163	-0.104	0.293
SD	0.937	0.903	0.933	0.905	0.979	0.872

Notes: The table shows treatment effects on summary indices of labor supply, social interactions, and preventive measures for the overall sample and the female subsample, respectively. Results are reported for individuals with high and low levels of religiosity, respectively. Significance levels are indicated by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 2.B.11: Effects on Labor Supply, Social Interactions, and Preventive Measures by Religiosity: Overall Sample

	Less Labor Supply		Less Social Interactions		More Preventive Measures	
	High Religiosity (1)	Low Religiosity (2)	High Religiosity (3)	Low Religiosity (4)	High Religiosity (5)	Low Religiosity (6)
Phone	0.017	0.003	0.014	0.006	0.001	0.022
	(0.038)	(0.029)	(0.020)	(0.016)	(0.022)	(0.018)
Phone + loudsp.	0.079	0.039	-0.027	0.014	0.006	0.036
	(0.058)	(0.038)	(0.034)	(0.022)	(0.033)	(0.027)
p-val phone=phone+loudsp.	0.275	0.324	0.224	0.706	0.872	0.589
p-val any treatment	0.455	0.709	0.721	0.601	0.926	0.157
Obs.	2,126	4,071	2,126	4,071	2,126	4,071
Villages	593	828	593	828	593	828
R^2	0.391	0.357	0.397	0.409	0.433	0.437
Control mean	-0.182	0.088	-0.103	0.054	-0.014	0.008
SD	0.947	0.932	0.507	0.559	0.618	0.625

Notes: The table shows treatment effects on summary indices of labor supply, social interactions, and preventive measures for the overall sample. Results are reported for individuals with high and low levels of religiosity, respectively. Significance levels are indicated by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 2.B.12: Effects on Labor Supply, Social Interactions, and Preventive Measures by Religiosity: Female Subsample

	Less Labor Supply		Less Social Interactions		More Preventive Measures	
	High Religiosity (1)	Low Religiosity (2)	High Religiosity (3)	Low Religiosity (4)	High Religiosity (5)	Low Religiosity (6)
Phone	0.028	-0.018	0.009	0.002	0.069	0.012
	(0.079)	(0.036)	(0.043)	(0.022)	(0.051)	(0.026)
Phone + loudsp.	-0.045	0.070	-0.122	-0.008	0.022	0.030
	(0.111)	(0.052)	(0.074)	(0.031)	(0.066)	(0.037)
p-val phone=phone+loudsp.	0.529	0.064	0.075	0.733	0.507	0.591
p-val any treatment	0.832	0.996	0.749	0.991	0.194	0.549
Obs.	613	2,279	613	2,279	613	2,279
Villages	280	634	280	634	280	634
R^2	0.515	0.399	0.549	0.429	0.497	0.450
Control mean	-0.007	0.275	0.051	0.175	-0.068	0.028
SD	0.968	0.879	0.605	0.576	0.601	0.665

Notes: The table shows treatment effects on summary indices of labor supply, social interactions, and preventive measures for the female subsample. Results are reported for individuals with high and low levels of religiosity, respectively. Significance levels are indicated by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Figure 2.B.2: Effects on Single Outcomes of the Knowledge Index: Overall Sample and Female Subsample
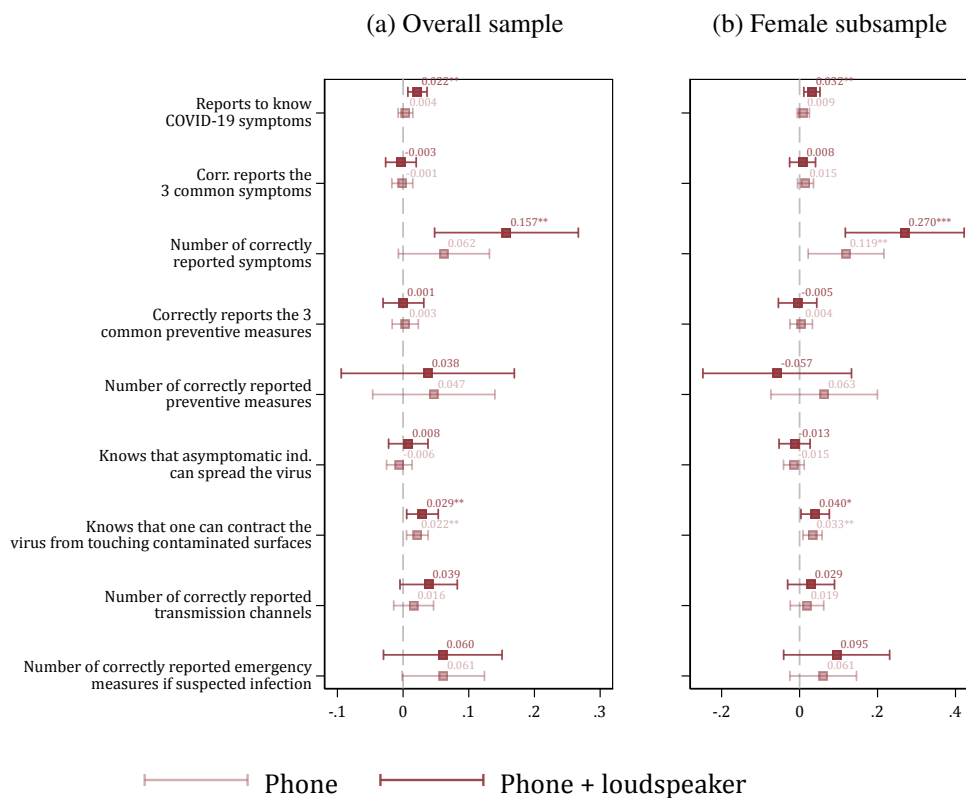
(a) Overall sample   (b) Female subsample

Phone   Phone + loudspeaker

Notes: The graph shows treatment effects on the components of the knowledge index for the overall sample (left) and the female subsample (right). Point estimates are shown with 10% confidence intervals. Significance levels are indicated by $* p < 0.10$, $** p < 0.05$, $*** p < 0.01$.

Figure 2.B.3: Effects on Labor Supply by Treatment Frequency
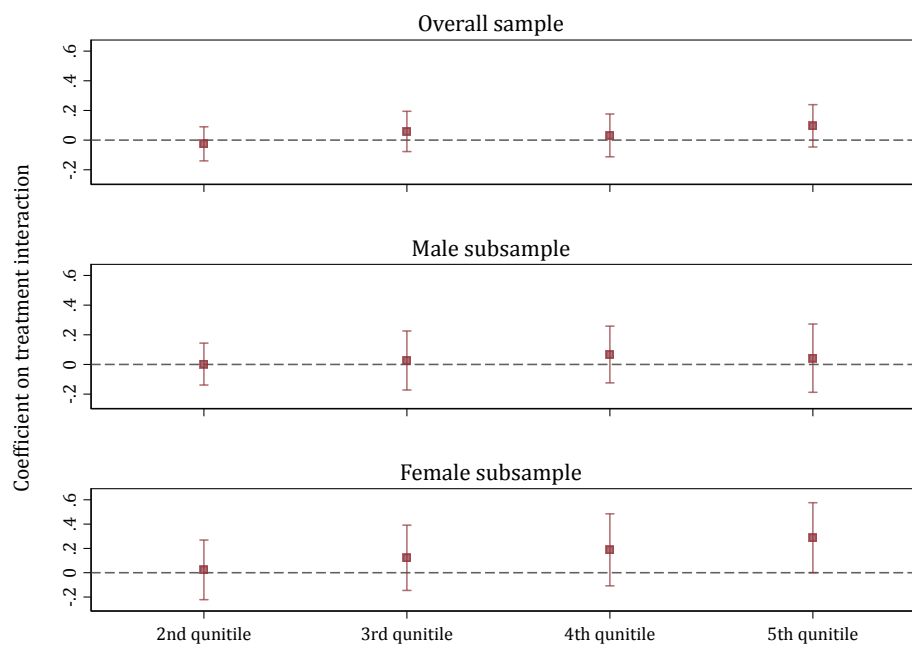
Chapter 3

The Economics of Women's Rights[*]

Joint work with Michèle Tertilt, Matthias Doepke, and Anne Hannusch

Abstract: Two centuries ago, in most countries around the world, women were unable to vote, had no say over their own children or property, and could not obtain a divorce. Women have gradually gained rights in many areas of life, and this legal expansion has been closely intertwined with economic development. We aim to understand the drivers behind these reforms. To this end, we distinguish between four types of women's rights—economic, political, labor, and body—and document their evolution over the past 50 years across countries. We summarize the political-economy mechanisms that link economic development to changes in women's rights and show empirically that these mechanisms account for a large share of the variation in women's rights across countries and over time.

3.1 Introduction

Two hundred years ago, a state of extreme inequality between the legal rights of women and men was the norm in most countries around the world. Political rights, where they existed, were usually reserved to men. In common law countries, such as the United States and the United Kingdom, married women had no say over their own property or children, and no protection against domestic violence; the husband exclusively exerted all rights of the married couple. Elsewhere, fathers' control over their daughters extended into adulthood and lasted until marriage, when another man would gain legal control.

Figure 3.1: Women's Rights and Economic Development across Countries
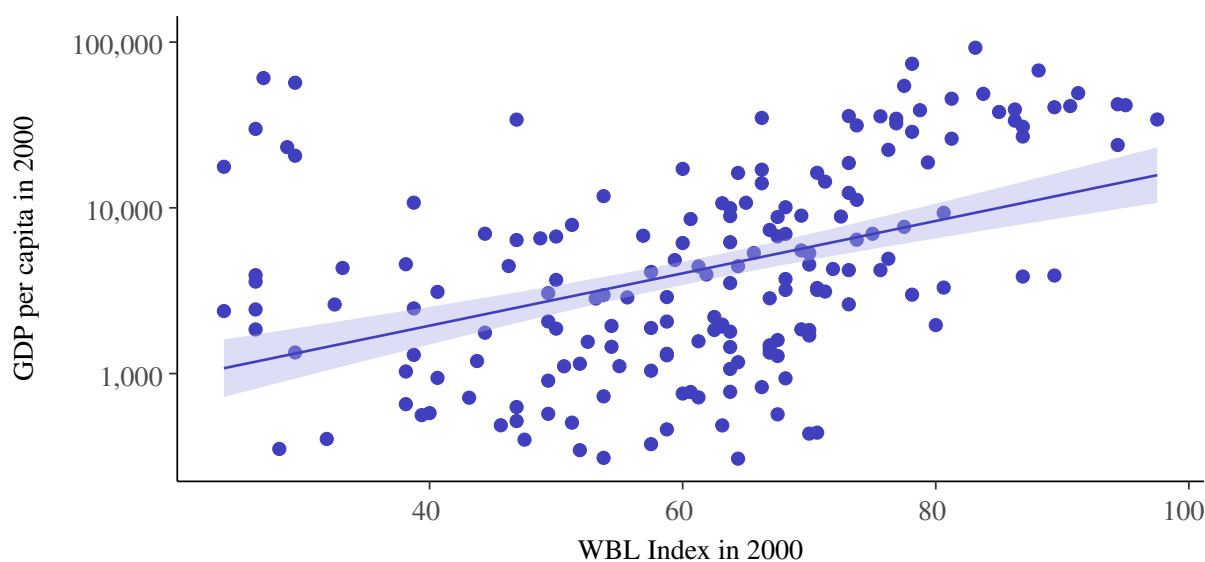
Notes: We plot the Women, Business and the Law (WBL) Index (World Bank, 2022a) as a measure of female empowerment against GDP per capita in constant 2015 US Dollars, where GDP per capita is on a log scale. The data for GDP per capita comes from the World Development Indicator database (see Appendix Table A.2). We include a linear regression of the WBL Index on log GDP per capita and show 90% confidence intervals.

In the time since, women's legal position has greatly improved in most countries. The starting point for this paper is the observation that expansions in women's rights have been highly correlated with overall economic development. Figure 3.1 plots an index of women's legal rights against GDP per capita in 2000, displaying a strong positive correlation between women's rights and development. The countries where women have the most rights (an index close to 100) have an income per capita above USD 10,000, while places where women have few rights (an index below 30) tend to be poor. A few exceptions of rich countries with a low women's rights score stand out; these are largely nations that owe their wealth to natural resources such as oil, and thus have become rich without having gone through the usual development process.[1]

---

[1] The six countries in the upper left corner of Figure 3.1 are all located in the Middle East: Bahrain, Kuwait, Oman, Qatar, Saudi Arabia, and the United Arab Emirates.

Despite much progress over the past two centuries, full equality between the sexes has yet to be achieved. The literature documents persistent gender gaps in a variety of outcomes such as education, employment, earnings, and wages (Goldin et al. 2006; Olivetti and Petrongolo 2016; Blau and Kahn 2017). Rather than outcomes, we focus here on gender gaps in formal legislation (de jure rights).[2] In most countries, women still do not possess the same legal rights as men. As shown in Figure 3.1, this gap is more pronounced in poor countries, though it also appears in rich countries. Table 3.1 provides several examples based on data from the year 2000. While women have the right to be elected in 97% of countries, there remain exceptions in certain high-income countries.[3] In 96% of all countries, women can sign contracts in the same way as men, but this is true in only 89% of low-income countries. In many places (a quarter of all countries), women still need their husband's consent to obtain a passport; this being a higher fraction for the low-income countries. Women have the right to work the night shift in the same way as men in about 70% of all countries, a percentage that drops notably among the low-income countries. Women can work in industrial jobs in the same way as men in only 41% of all countries, and in only 23% of low-income countries. Legal protection against domestic violence exists in even fewer countries and not at all in low-income countries.

Table 3.1: Examples of Women's Rights by Income

| | % of Countries by Income | | |
Right to …	All	Low	High
… be elected	96.7	100.0	88.9
… sign a contract in the same ways as a man	95.8	88.5	100.0
… obtain a passport in the same ways as a man	75.3	69.2	82.5
… work night shift the same way as a man	68.9	61.5	80.7
… work in an industrial job the same way as a man	41.1	23.1	45.6
… legal protection against domestic violence	24.2	0.0	45.6

Notes: We report the percentage of countries in which women have a certain right. The statistics are based on the year 2000. We use individual questions about women's rights and country income classifications from the Women, Business and the Law Database (World Bank, 2022a). Information on the right of women to be elected comes from Paxton et al. (2008).

The expansion of women's rights has unfolded gradually over time. Starting in the nineteenth century, women began to gain basic economic rights in most of today's high-income countries, such as the right to own property. Political rights, most importantly the right to vote, came next, introduced in many high-income countries in the first half of the twentieth century.

---

[2] Due to a lack of enforcement, legislation does not always translate directly into de facto rights, though it does provide an important first step in guaranteeing equal treatment of women and men.

[3] Namely, the same oil-rich countries mentioned in footnote 1.

More recently, laws relating to equality in the labor market have come into existence.[4] Still in process are women's rights over their own bodies, such as protections from domestic violence and marital rape.

The goal of this paper is to understand the political economy of women's rights. What causes reforms? Is there a direct link between economic development and expansions of women's rights, or is the correlation between the two accidental? To what extent can economic analyses contribute to understanding these issues?

Understanding the drivers of reforms necessitates study of the incentives of individual voters or politicians to support them. Doing so sheds light on why certain rights were introduced at particular points in time, why they correlate with economic changes, and why women in some countries still lack many rights. While there is a sizeable literature on the consequences of expanding rights to women, much less research has investigated the origins of these rights. In addition to elucidating the drivers of reforms and specifically the role of economic forces, we aim to identify the main open questions that beg further attention.

In political-economy models of reform, political preferences are based on individual preferences: Voters and politicians favor or oppose particular reforms based on how the reforms would change outcomes that they or their family members care about. A possible alternative view is that the extension of rights to women is simply another incarnation of a broad expansion of rights—first from the elites to the masses, from the wealthy to poor, and later to various minorities and even animals—driven by other forces like changes in culture and religious doctrine or a general enlightenment and better awareness of the needs of others. In part, our goal here is to assess just how powerful economic theories of political change are relative to such alternatives. One argument in favor of a major role for economic mechanisms in expansions of women's rights is that women are linked to men in ways that other groups are not. Men are fathers, husbands, and sons of women and are thus directly and economically connected to them. In contrast, many people may never have direct contact with members of groups whose rights are at stake in other reforms.

We begin this paper with a historical description of the expansion of women's rights in the United States. By highlighting several influential laws and landmark court decisions, we illustrate how the focus of legal change has shifted over time, first from economic to political rights, then to labor market rights, and most recently to body rights.

We subsequently present a theoretical framework that highlights four important economic channels for political-economy models of changes in women's rights. The *bargaining power channel* concerns the effect that rights have on the sharing of resources within any given family, implying that women should generally be in favor and men against women's rights. The

---

[4] See also Hyland et al. (2020), who document women's rights relative to the labor market over the period of 1970–2020 in 190 countries.

parental altruism channel refers to the idea that fathers care about their children and women's rights would thus benefit men directly through the effect on their daughters, but also indirectly, as women's rights may increase investments in children. The *income channel* reflects the notion that more women's rights can in some circumstances increase total resources (for example, if an agency problem is solved), which can in turn benefit men. Finally, the *public policy channel* relates to the idea that including women in policy-making can alter the chosen policies. We explain the importance of these four channels in various episodes of expansion in women's rights in the United States. In doing so, we also discuss the existing theories on the topic.

We then conduct an empirical analysis of women's rights across countries during the more recent period of 1970-2021. We construct different indices of women's rights related to the four areas: economics, politics, labor, and body. Using cross-country panel data for 190 countries, we regress each rights indicator on several economic and cultural variables. We find that overall economic development, as proxied by GDP per capita, is positively correlated with each of these aspects of women's rights.[5] However, the correlation with GDP per capita becomes much smaller or disappears once we control for specific measures—namely the TFR and women's labor force participation—that proxy for the parental altruism and income channels for changes in women's rights. These specific variables correlate with all aspects of women's rights and account for a substantial share of cross-country variation, and are particularly relevant for economic and labor rights. Yet, even for body rights, a considerable portion of cross-country variation can be accounted for by differences in economic variables, with the fertility rate playing the largest role. To explore the role of cultural factors that may work independently of economic channels, we alternatively include country fixed effects and control for religion variables. In these specifications, the economic channels continue to display a robust correlation with women's rights. Though religion variables appear to play a more limited role, in some cases they diminish the positive impact of economic development on legal reform. Overall, at least at the level of correlations, there is strong support for economic mechanisms, which, perhaps unsurprisingly, are particularly powerful in accounting for variation in economic and labor rights.

Our results suggest fruitful directions for future research. While our cross-country empirical analysis is suggestive, clearly there is much scope for work that more directly identifies the impact of specific channels on legal reform. On the modeling side, the majority of the existing literature focuses on historical reforms in high-income countries, but few studies use formal models of political change to address variation in women's rights in the cross-section of

---

[5]   In principle, of course, the causality could also run in the opposite direction. While evidence on specific rights leading to more investments and hence higher incomes exists (see Duflo 2012, Doepke et al. 2012 and Doepke and Tertilt 2016 for discussions), it is unlikely that much of the income growth since industrialization can be attributed to the expansion of women's rights.

countries today, including low-income countries.

3.1.1 Related Literature

Doepke et al. (2012) provide a survey of earlier economic literature on women's rights that focuses on the mutual interaction of economic and legal changes. Theories on the expansion of women's rights have largely focused on economic rights in the United States. Geddes and Lueck (2002) relate the expansion of women's economic rights to women's role in the labor market, while Doepke and Tertilt (2009) and Fernández (2014) develop theories relating these same rights to women's role in educating children. A number of papers empirically analyze the consequences of the expansion of women's economic rights in the United States (Khan 1996; Geddes et al. 2012; Alshaikhmubarak et al. 2019; Hazan et al. 2019; Hazan et al. 2021).

Bertocchi (2011) develops a theory of the extension of female suffrage. Jones (1991) and Braun and Kvasnicka (2013) empirically examine the reasons behind the expansion of suffrage in the United States, while Teele (2018) documents the role of the suffrage movement in driving reform in the United States, France, and the United Kingdom. Other papers assess the impact of extending suffrage to women on outcomes such as government spending in the United States (Lott and Kenny 1999; Miller 2008), Europe (Aidt and Dallal, 2008), and Switzerland (Abrams and Settle 1999; Funk and Gathmann 2014; Slotwinski and Stutzer 2022).

Theories of women's expanding political rights also relate to an economic literature on the general spread of political rights. As early as 1959, Lipset (1959) argued that economic development is a prerequisite for democracy. The hypothesis that development leads to democratization has been confirmed empirically (Barro 1999; Murtin and Wacziarg 2014) and modelled theoretically. Acemoglu and Robinson (2000) and Lizzeri and Persico (2004), for example, build political-economy models of the expansion of rights from the elites to the masses.

A large number of studies analyze the impact of new laws related to women's rights in the labor market (Landes 1980; Zabalza and Tzannatos 1985; Goldin 1988b; Marchingiglio and Poyker 2019; Bailey et al. 2021), though few examine the origins of these laws. In an early contribution, Huber (1976) discusses the importance of technical change for the women's movement, specifically in relationship with labor laws. Hunt and Rubin (1980) argue that labor market rights are related to the number of single women in the economy who have the most to gain from such legislation and confirm this hypothesis empirically in U.S. cross-state data. Goldin (1988a) finds that marriage bars were associated with modern personnel practices and argues that they disappeared when the cost of limiting female labor supply became too high in the 1950s. In a previous Marshall Lecture, Pande and Roy (2021) document a strong correlation between labor law equality and social norms about working women. To our knowledge, Doepke et al. (2021) offer the only formal theory of the political economy of labor laws affect-

ing women.

Work on women's rights in other fields such as sociology and political science has often emphasized non-economic forces. For example, Htun and Weldon (2018) highlight the role of feminist movements in bringing about change in dimensions such as workplace equality. They also argue that a changing relationship between religious bodies and the state underlies reform in areas such as family law, which used to be regulated by religious institutions. Most studies on the origins of female body rights have been conducted outside of economics. Within sociology, Boyle et al. (2015) assess the liberalization of abortion policies throughout the late 20th century around the world. The authors find that the influence of modern science and medicine and the number of women in parliament have contributed to liberalizing abortion policies. Catholicism, on the other hand, is negatively related to abortion liberalization. Ebetürk (2021) looks specifically at child marriage,[6] and finds that female legislators are an important driver of bans on this practice while religion (in this case, Islam) slows down reforms. Other scholars examine the consequences of body law reforms. Frank et al. (2009) look at rape law reforms. Godefroy (2019) empirically analyzes a reduction in women's rights related to sexual behavior in Nigeria, while Tertilt (2006) theoretically explores the impact of giving women property rights over their own body in a polygynous society.

Changes in women's legal rights are related to shifts in social norms about gender roles.[7] The relationship between changes in social norms and those in legal rights is complex. On the one hand, legal changes shape expectations and can hence modify social norms over time. On the other hand, shifts in individual attitudes must necessarily precede legal changes since a political majority is needed for reform to emerge. We argue that it is often changing economic conditions that drive both changes in social norms and, eventually, in formal rights. For example, following Boserup (1970), Alesina et al. (2013) relate traditional gender norms to plough agriculture, while Becker (2019) connects the practice of female genital cutting to the desire to reduce paternity uncertainty in pastoral societies. While some of these practices and norms persist through time despite the disappearance of their original cause, many do not. Cheung (1972) relates the practice of footbinding in China to the desire to establish property rights over girls and wives and Bossen et al. (2011) argue that the practice disappeared when the arrival of commercial cloths made weaving and spinning at home unprofitable and it was no longer necessary to keep girls at home. Fernández-Villaverde et al. (2014) suggest that changes in contraceptive technology affected social norms about premarital sex. In last year's Marshall lecture, Pande and Roy (2021) contend that the persistence of social norms prescribing separate

---

[6] While laws against child marriage in theory equally affect men and women, in practice they mostly protect teenage girls from being married without their consent, and are thus considered here as an important women's body right.

[7] See Jayachandran (2015) and Giuliano (2022) for two surveys, the former discussing the roots of gender inequality specifically in developing countries, the latter focusing on the historical origins of gender norms.

spheres for women and men is, at least in part, connected to the rents associated with preferential access to well-paying jobs. Social norms are not simply the sum of individual attitudes, rather they are beliefs about perceived attitudes of others. To this regard, Bursztyn et al. (2020) demonstrate the effect that misperceptions of others' attitudes can have. In a field experiment in Saudi Arabia, the authors find that more women start to work in a for-pay job if informed that the true attitudes in society are actually more favorable towards working women than they had previously thought.

3.2   Expansion of Women's Rights in the United States

Today, men and women in the United States are close to equal in terms of legal rights. While substantial gender differences remain in outcomes (e.g., lower labor force participation for women and a sizeable gender wage gap), few of these differences originate from unequal rights. This has not always been the case, particularly for married women. Until the early 19th century, women lost their separate legal identity upon getting married, when the legal rights of husband and wife were merged and subsequently exercised solely by the husband. As a consequence, married women could not sign a contract, own property, or decide how to spend family money. They typically could not initiate divorce or gain child custody in the event of a separation. The legal position of single women was somewhat better, though they still had fewer rights than men (e.g., they could not vote).

The legal position of American women began to change in the mid-19th century. As illustrated in Figure 3.2, the expansion of women's rights occurred in four main phases, each revolving around a distinct class of rights. First, over the second half of the 19th century, women obtained basic economic rights. Second, in the early 20th century, women were granted political rights. Much later, in the 1960s and 1970s, laws regulating the legal equality of women in the labor market were passed. Finally, starting in the 1970s and still ongoing, women's rights related to their own bodies have substantially improved.
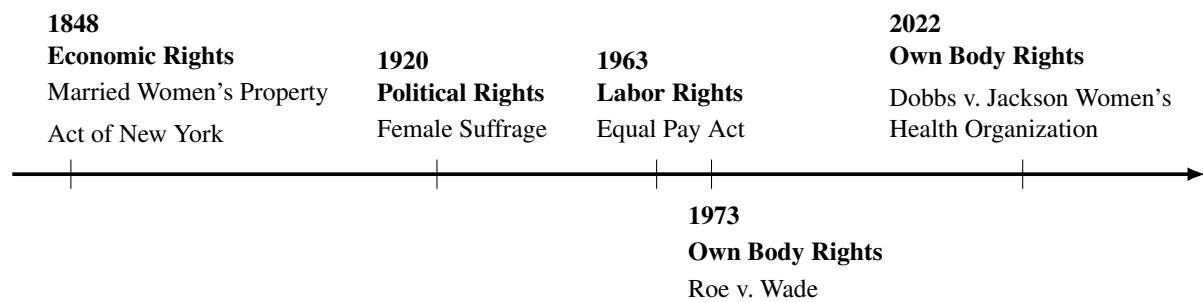
To illustrate the gradual expansion of women's rights, in what follows we describe a series of key law changes reflective of each successive phase of legal reform. This is not meant to be exhaustive, rather the aim is to underline the fact that distinct types of women's rights were passed at different points in time throughout U.S. history.[8]

The first type of rights women obtained consisted of basic economic rights, such as that to hold property.[9] In 1839, Mississippi was the first state to grant married women limited

---

[8] For more comprehensive reviews of the U.S. history of women's rights, see Hecker (1971), Salmon (1986), Yalom (2001), and the timeline in Doepke et al. (2012). More recent developments are discussed in McBride and Parr (2010). For abortion law specifically, see Baker (2022).
[9] Geddes and Tennyson (2013) provide excellent data on the passage of married women's property and earnings

Figure 3.2: Influential Laws and Landmark Rulings that Impacted Women's Rights

1848
Economic Rights
Married Women's Property
Act of New York

1920
Political Rights
Female Suffrage

1963
Labor Rights
Equal Pay Act

2022
Own Body Rights
Dobbs v. Jackson Women's
Health Organization

1973
Own Body Rights
Roe v. Wade

property rights. The more comprehensive Married Women's Property Act of New York was passed in 1848, and quickly became the model to which other states looked. Additional legal developments during the period of 1850-1900 included earnings laws, sole trader laws, child custody reforms, and the right to initiate a divorce. All these rights were extended to women at a time when only men could vote.

In a second phase, women gained political rights. A few states, mostly in the West, had already passed state suffrage laws in the late 19th century. However, at the federal level, universal female suffrage was not introduced until 1920, when the Nineteenth Amendment to the U.S. Constitution came into effect.

During the third phase (i.e., the 1960s and 70s), important law changes granted women equality in the labor market. Previously, "marriage bars" could legally exclude married women from working in certain professions. These were widely applied, a notable example being school teachers: in most states, a woman school teacher had to resign upon marriage. In fact, during the first half of the 20th century, many new laws were introduced that put restrictions specifically on female workers: on night work, maximum hours, minimum wage, and even on seating, which required employers to provide a chair for each female employee. These asymmetries between male and female workers came to an end in the 1960s. With the Equal Pay Act of 1963 it became illegal for employers to differently compensate women and men for the same work and Title VII of the Civil Rights Act of 1964 eliminated most other labor market asymmetries between female and male employees. The 1960s and 1970s were also characterized by debates over the Equal Rights Amendment, which gained increasing support with the rise of the women's movement in the United States. The amendment proposed a change to the U.S. Constitution that would guarantee equal rights to U.S. citizens independent of their sex and was first introduced in the 1920s. While passed by the senate in 1972, it was never ratified nationally. Eventually, however, roughly half of the states passed state-level Equal Rights Amendments (Wheaton, 2021).

---

acts across U.S. states.

Lastly, beginning in the 1970s and ongoing, a series of laws related to women's own bodies have been passed.[10] In 1973, the U.S. Supreme Court ruled in the landmark *Roe v Wade* case that the right to have an abortion was a fundamental right to privacy and thus a constitutional right. More than fifty years later, in 2022, the Supreme Court overturned this ruling in *Dobbs v Jackson Women's Health Organization*, and several states subsequently made abortion illegal. This example reflects the fact that women's rights have not always progressed in a linear fashion: reversals do happen.[11] Certainly, the right to physical integrity is not limited to the issue of abortion and includes many other dimensions. Throughout the late 20th century, many laws were introduced that made marital rape, teenage marriage, domestic violence, and sexual harassment all illegal and punishable by law.

3.3   Economic Theories for the Expansion of Women's Rights

What economic theories might best help explain the spread, and occasional reversal, of women's rights? By economic theories, we mean explanations that employ an economic model of human behavior, wherein people's political preferences are based on their individual preferences over various outcomes that directly affect them. Importantly, these outcomes may extend beyond a narrower delimitation of economic variables such as income, wages, and prices, to also include concepts such as altruistic concern for one's children. Broadly, an economic explanation for reforms in women's rights would argue that women's rights change individuals' constraints, choice sets, and outcomes. These potential changes, in turn, determine political preferences, which through a political mechanism (e.g., voting) then generate a political outcome.

We view such economic explanations as complementing and competing with alternative explanations that are not based on the effects of women's rights on individual outcomes. For example, a religious belief of "what is right" independent of one's own life would fall into this category; so too would a "spread of enlightenment" view where people, over time, change their abstract conceptions of who does or does not deserve certain rights, again independently of their personal lives. Of course, it is possible to represent such explanations using the economic tool set; for example, by hard-wiring religious or other preferences into the utility function. The key distinction here, however, is that what we term economic explanations hinge on the effects that women's rights have on the individual outcomes that people care about.

Put differently, we aim to identify who wins and who loses from the introduction of specific

---

[10] While one could argue that female body rights are distinct from other equality laws in that they apply only to women, we contend that such laws guarantee equality in protecting physical integrity while recognizing physical differences between the sexes.

[11] Brooks et al. (2022) argue that the abortion right reversal is specific to the U.S. context and will not halt reforms in other countries.

rights, based on their individual preferences. In this view, women's rights will be introduced if there is a winning coalition that stands to gain from these rights—our task is to identify this coalition and what motivates them. Political change can occur when economic shifts (such as an increase in the return to human capital) alter the political tradeoff faced by pivotal groups, or alternatively when there are changes in the size of groups who favor and oppose reform (such as an increase in the number of single versus married individuals).

To build such an economic model of the political economy of women's rights, we must start by specifying people's preferences. Here, we present a simple setup to illustrate the main mechanisms that have been used in the literature. In this setup, the lifetime utility of an individual of gender $g \in \{f, m\}$ can be represented as:

$$V_g(h_f, h_m, X) = u_g(c_g, 1 - n_g, P) + \gamma_g \, V_C(h_C, X'). \tag{3.1}$$

V_g denotes total lifetime utility as a function of the individual state variables h_f, h_m (human capital of wife and husband) and the aggregate state variable X, which captures the current legal regime. For a single individual, only own human capital would enter as a state variable. The period utility function depends on individual consumption c_g, leisure $1 - n_g$ (n_g is labor supply), and an aggregate public policy variable P. People also care about the welfare of their children; γ_G is accordingly the degree of parental altruism, and $V_C(h_C, X')$ is the lifetime utility of one's children, which depends on their human capital h_C and the future legal regime X' to which the children will be subject.

Starting from this specification of preferences, we now illustrate the main channels that have been used in the literature to link women's rights to individual preferences and, ultimately, political preferences. We call these the *bargaining power channel*, the *parental altruism channel*, the *income channel*, and the *public policy channel*.

3.3.1 Bargaining Power Channel

The most direct effect of expansions of women's rights is the broadening of women's opportunities, which may come at least in part at the expense of men's opportunities. From this perspective, women should be in favor of expanding women's rights, and men should be against. A common way of modeling such a conflict between women and men is to allow for a *bargaining power channel*, whereby women's rights help determine the allocation of resources within marriage. To illustrate this channel, consider a population of married couples with preferences

given by (3.1) where the determination of consumption c_f and c_m can be written as follows:

$$c_f = S_f(h_f, h_m, X)I(h_f, h_m),$$
$$c_m = (1 - S_f(h_f, h_m, X))I(h_f, h_m).$$

Here $I(h_f, h_m)$ represents the income earned by a couple with human capital h_f, h_m, and $S_f(h_f, h_m, X)$ is the share of consumption going to the wife given the individual state variables and the political regime X. If a larger X represents more rights for women and if women's share in marital consumption is increasing in rights, $\partial S_f(h_f, h_m, X)/\partial X > 0$, women are going to benefit from expansions of women's rights, and men are going to lose. Such a relationship from rights to bargaining power can be formalized, for instance through marital bargaining subject to outside options that depend in part on legal rights, like those related to divorce and marital property laws. Models along these lines have been widely used in the family economics literature, going back at least to Manser and Brown (1980). Voena (2015) and Foerster (2021), for example, analyze how laws specifying the division of property and alimony payments upon divorce affect consumption and welfare of spouses in marriage. These models have found empirical support (Stevenson and Wolfers 2006; Mazzocco 2007; Lise and Yamada 2019).

The strength of the bargaining power channel may depend on the specific right at stake. It is, for example, more likely to be important for laws that specifically address the rights of married women; although laws that shift, say, the labor market opportunities of single women could still matter by shaping outside options. Moreover, there is heterogeneity in people's exposure to the bargaining power channel. For instance, this channel will matter little for singles who expect to remain single for a long time or permanently. A change in household composition, such as the increase in the share of single households observed in many countries, could therefore influence the role of this channel.

A bargaining power channel is present in a number of formal studies of expansions of women's rights, though it is usually not the only channel. Indeed, under this channel, only women would support women's rights, which would contradict the fact that many women's rights were introduced or expanded before women had the right to vote. More generally, political support for different forms of women's rights does not usually divide sharply along gender lines. The literature has therefore identified additional channels where support or opposition to rights expansions does not necessarily depend solely on one's own gender. As according to the bargaining channel men only stand to lose from introducing women's rights, at least one of these additional channels must be operative for initial reforms to happen.

3.3.2 Parental Altruism Channel

One such additional channel is what we term the *parental altruism channel*. Parents care about their children, and reforms to women's rights can have repercussions for daughters and sons that, in turn, translate into political preferences. In the utility function (3.1), parental altruism is represented through the children's utility $V_C(h_C, X')$, which enters parental utility with weight γ_g. The children's utility is an average of both daughters' and sons' future utility; if there are equal numbers of daughters and sons, we have:

$$V_C(h_C, X') = \frac{1}{2} \left(V_f(h'_f, X') + V_m(h'_m, X') \right).$$

If women's rights solely had distributional implications (i.e., making women better off and men worse off), parental altruism would already imply a preference for some degree of women's rights, because with curvature in utility parents would like to reduce inequality between their sons and daughters. This motive for supporting women's rights is further strengthened if women's rights also affect the human capital of children. It is often argued that women are more altruistic towards children than are men, $\gamma_f > \gamma_m$.[12] In models of household bargaining, this implies that children's human capital is increasing in women's bargaining power, and accordingly in the extent of women's rights X:

$$h_C = F(h_f, h_m, X),$$

with $\partial F(h_f, h_m, X)/\partial X > 0$.

The parental altruism channel lies at the heart of the model proposed by Doepke and Tertilt (2009), which analyzes the introduction of women's economic rights in the nineteenth century. As described in Section 3.2, a notable feature of the reforms to marital property, divorce, and child custody laws in both the United States and the United Kingdom during this period is that they were introduced long before women gained the right to vote. A political-economy explanation of these changes must therefore necessarily focus on the preferences of men, who as voters and politicians brought about these changes. In the model of Doepke and Tertilt (2009), the bargaining power channel pushes men to oppose women's rights, as they imply a loss of bargaining power in their own marriage. Simultaneously, however, the parental altruism channel provides a rationale to support women's rights. The political outcome hinges on the relative strength of the two channels. The power of the parental altruism channel depends crucially on the importance of human capital investment for children's future welfare. The authors argue that, until

---

[12] This contention can be supported by arguments from evolutionary biology: men face a higher paternity uncertainty than women, while women are more constrained in their reproductive capacity. There is a sizeable empirical literature that suggests that women place more importance on children than do men, although this evidence is not unambiguous (see Doepke and Tertilt 2019).

the early nineteenth century, human capital was of limited importance and support for women's rights consequently remained low. Subsequently, however, the demand for human capital grew substantially, leading to the well-documented advent of mass education during the second half of the nineteenth century. Doepke and Tertilt show that this same technological shift increased the strength of the parental altruism channel, ultimately leading to political reforms. In line with the tradeoff between these two channels, they document that during this historical period, political debate over the introduction of women's rights focused on the tradeoff between the rights of husbands and the implications of unchecked rights for men for the welfare of children. Thus, it was ultimately the transition from an agricultural society to a modern knowledge-based economy that led to the expansion of economic rights at a time when women had few other rights.

Fernández (2014) similarly develops a model that explores the implications of a tradeoff between the bargaining power channel and the parental altruism channel for the introduction of women's economic rights. She looks specifically at men's desire to be able to transfer bequests to their daughters, which requires women to have property rights. In a poor economy in which there is little capital and bequests are consequently small, this motive is not strong enough to outweigh the bargaining power channel. However, in a growing economy, the parental altruism channel becomes more important as capital stock rises and fertility rates fall, and ultimately men agree to expand women's rights. Fernández finds support for some predictions of this theory in cross-state data in the United States. Hazan et al. (2021) provide further empirical support for the theories advanced by Doepke and Tertilt (2009) and Fernández (2014), showing that the expansion of married women's property rights in the United States led to an increase in education and a decline in fertility.

A general implication of the parental altruism channel is that support for women's rights should be increasing in one's number of daughters. Washington (2008) finds empirical support to this regard, showing that politicians with more daughters are more likely to vote liberally on women's issues and specifically support reproductive rights. Oswald and Powdthavee (2010) observes that having daughters makes voters more likely to support left-wing parties. Conversely, in the aggregate the parental altruism channel should be less powerful if there are many childless individuals.

3.3.3   Income Channel

The models of Doepke and Tertilt (2009) and Fernández (2014) focus on intra-household bargaining and on investments in children, ignoring women's participation in the labor market. This is because in the nineteenth and well into the twentieth century, married women's labor force participation rates were quite low, suggesting that it was largely other factors that drove

reforms in married women's rights during this period. In more recent times, characterized by much higher female labor force participation, women's work is essential to laws affecting the labor market and, more generally, for reforms. This shift motivates the income channel, which hinges on general-equilibrium implications of women's labor force participation.

Consider a population with preferences given by (3.1) where consumption for single individuals of gender $g \in \{f, m\}$ is given by $I(h_g, X)$ and for married individuals we have:

$$c_f = S_f(h_f, h_m, X)I(h_f, h_m, X)$$
$$c_m = (1 - S_f(h_f, h_m, X))I(h_f, h_m, X).$$

As in Section 3.3.1, we allow that the legal regime X affects bargaining power within marriage. The new angle is that now income $I(h_g, X)$ for singles and $I(h_f, h_m, X)$ for couples also depends on women's rights X. The focus here is on rights that affect women's ability to work, either by directly influencing the regulation of women's labor (such as marriage bars, overtime restrictions, and occupational limitations) or by affecting their incentive to work (for example, through giving them control over their own earnings). The dependency of income on X captures two different mechanisms. First, there is a direct impact of changes in rights X on the income of women who respond to the change by working more or by working for higher wages. If women's labor rights are expanded, this direct effect is generally positive and will lead women (and their husbands) who directly benefit from the legal change to support reform.

A second, indirect effect occurs because the expansion of women's labor supply through the direct effect changes the general equilibrium wage structure. Existing workers of either gender who compete with women who would enter after reforms in the labor market would see a reduction in earnings, as the expansion in the supply of their skill type drives down relative wages. Conversely, workers whose labor supply is complementary to that of entering women would see higher wages.

Broadly, support for the expansion of women's rights based on the income channel depends on two factors: (i) the substitutability of one's own labor with that of women who would enter the labor market and (ii) one's own initial labor supply and that of one's spouse. Support for or opposition to women's rights will thus not primarily depend on one's own gender. Women who are already working may be opposed to more women entering, as this would lower their wages, just like unions often oppose more competition for their own workers. Likewise, women who already decided not to work no matter what (say, to focus on raising their children) would oppose expanded labor rights for women if their husbands compete with women in the labor market. Conversely, men whose labor is complementary to that of women and married men whose spouse would benefit from higher earnings opportunities would support reform.

The income channel is central to the analysis of Geddes and Lueck (2002), who study the

expansion of women's economic rights in the United States. They argue that, on the one hand, if women have no rights, husbands choose the time allocation of their wives but face an enforcement problem (the wife can shirk into leisure), which reduces women's labor supply and hence family income (the income channel). On the other hand, men get a larger share of household consumption (the bargaining power channel) when women have no rights. If there is an increase in the return to women's market work, the desire to raise family income out-weights the distributional motive and men start to support granting more rights to women. One implication of this theory is that women's economic rights should increase the incentive to invest in girls' human capital. Geddes et al. (2012) find support for this argument, showing that the expansion of women's economic rights increased girl's school attendance relative to boys. The income channel is also related to work by Khan (1996) who finds that greater property rights stimulated female patenting and commercial activity and Hazan et al. (2019) who show that those same rights led households to shift their portfolios towards financial assets and more rapid industrialization. The authors argue that the reason men extended these rights to women was precisely because it would increase overall income through an improved allocation of capital.

Doepke et al. (2021) use a political-economy model to better understand the rise of regulations that limited women's labor rights in the United States (including overtime restrictions and marriage bars) as well as their abandonment later on. When restrictions were introduced between 1880 and 1940, women entered the labor force in larger numbers, but relatively few married women were working. The authors argue that as a result, there was a broad coalition in favor of restricting women's work that included single men competing with women in the labor market, married men whose own wives were not working, and non-working wives who were concerned about competition for their husbands. The coalition in favor of restricting women's labor began to shrink when rising returns to working drew more married women into the labor market, which implied that both these married women and their husbands stood to gain from removing restrictions. Therefore, the income channel is key for understanding reforms to labor rights in the mid-20th century. Hunt and Rubin (1980) argue instead that the increase in the number of single women played an important role in the abandoning of such discrimination as they stood most to gain. They find some support for this argument in U.S. cross-state data. Of course, both forces may have operated simultaneously to create a coalition of single women and dual-earner couples in favor of labor market reforms.

3.3.4   Public Policy Channel

The channels analyzed thus far concern legal changes that have a direct impact on people's economic lives, for example by changing property or labor law. The introduction of women's suffrage differs in that this reform did not impose any immediate modification of the rules and

constraints that affect families and individuals. Rather, it changed future political outcomes by altering the makeup of the electorate. If women's political preferences and voting behavior were identical to men, this would have no direct effects at all, at least from the perspective of our choice-based economic analysis. In our framework, political preferences over women's political rights must therefore necessarily derive from differences in political preferences between women and men and what these imply for future political outcomes.

In the utility function (3.1), general preferences over political outcomes are captured by the argument P in the period utility function $u_g(c_g, 1-n_g, P)$. Extending political rights to women, once again denoted by the political state variable X, changes the composition of the electorate and hence the process that determines P, so that we can write $P = F(X)$. Political preferences for extending suffrage to women then depend simply on how aligned a given voter's political preferences are with women's political preferences. If we let $X = 1$ denote women's suffrage and $X = 0$ the regime where only men vote, an individual will be in favor of women's suffrage if they prefer the public policy $\tilde{P} = F(1)$ over the status-quo policy $P = F(0)$.[13]

Once again, political preferences for extending rights to women may not necessarily depend on gender. In the polar case where all men are identical, they would not support female suffrage. When there is variation in political preferences, men whose preferences are more aligned with those of women will support their right to vote. Conversely, women whose preferences are more in line with those of men than with those of other women would oppose female suffrage.

To provide content to the public policy channel, it is necessary to know how women's and men's political preferences differ in the data. A sizeable literature documents the effects of female representation on government spending. Lott and Kenny (1999), for example, show that female suffrage in the United States increased government spending, while Miller (2008) more specifically documents that this reform raised public health spending and led to a decline in infant mortality. Chattopadhyay and Duflo (2004) find that female leadership translates to greater investments in infrastructure that is relevant to the needs of women. There is also empirical evidence that women's representation in government councils increased spending on childcare and education in Sweden (Svaleryd, 2009).

Bertocchi (2011) develops a formal theory of women's enfranchisement that builds on the public policy channel. In her model, there is a cost of keeping women disenfranchised, which could represent the discord between genders resulting from disenfranchisement or a cultural preference for more equality. At the same time, women have a higher preference for public-good spending and a higher preferred tax rate than men. As long as the gap in preferred policies between women and men is large, men withhold the right to vote from women. Over time, however, economic growth diminishes the relative return to physical strength (of which men

---

[13] There could be additional effects through the impact of X on children's utility; here, we focus on the case in which future policy X' is held fixed and only women's suffrage in the current period is in question.

have more), thereby lowering the gender wage gap, which in turn narrows the divide between women's and men's preferred policies. When the preferred tax rates of women and men are sufficiently close, the male median voter prefers to extend suffrage to women.

In line with Bertocchi's hypothesis that relative wages affect policy preferences, Edlund and Pande (2002) argue that changes in women's and men's relative income have driven recent political shifts in the United States In the early 1980s, a political gender gap began to emerge whereby more women than men favor the Democratic party. The authors link this shift to declining marriage and rising divorce rates, which they argue has made women relatively poorer and moved their policy preferences to the left.

3.3.5   Summary of Theories of Women's Rights

In the above sections, we described the main channels underlying changes in women's economic rights that have been explored within the broad framework of economic theories of political change. Although this literature has grown in recent years, the focus has largely been on specific historical episodes, leaving a number of open questions. Table 3.2 summarizes the topics that have to date been addressed and where gaps remain. Notably, many of the studies that include formal models of historical changes in women's rights refer to the United States and (to a lesser extent) other high-income countries such as the United Kingdom. Little to no work explores the ability of such models to account for heterogeneity in women's rights in the entire cross-section of countries. Many reforms have occurred in lower-income countries in recent decades, and it remains to be seen whether the existing theories can explain these shifts in low-income countries as well. There is also a lack of models that address changes in women's body rights, which include distinct aspects such as protection from domestic violence and abortion rights. This is despite the fact that the empirical literature documents wide-ranging implications of body rights for women, from future career prospects to the labor market more generally (see, for example, Goldin and Katz 2002; Bailey 2006; Aizer 2010; Myers 2017). Expanding formal economic modeling to such rights would be a promising direction for future research.

Beyond extending research to these underexplored areas, there are two other important directions future research might take. One consists of linking economic models of political change to empirical evidence on specific political reforms. The economic models provide clear predictions of who stands to gain and lose from women's rights, and how coalitions in favor of reform are shaped by other economic changes. Identifying more directly the source of political support for particular reforms will help distinguish between competing models and guide the development of new theories. A second area concerns the need to understand the relative importance of the broadly conceived economic motives described here and other forces such as cultural change or general enlightenment, and how these channels interact. As a starting point for such

Table 3.2: Existing Economic Theories about the Expansion of Women's Rights

Type of Right	Historical Reforms	Cross-Section of Countries Today
Economic	Geddes and Lueck (2002), Doepke and Tertilt (2009), Fernández (2014)	?
Political	Jones (1991), Bertocchi (2011), Braun and Kvasnicka (2013)	?
Labor	Hunt and Rubin (1980), Doepke et al. (2021)	?
Own Body	?	Papers in sociology, e.g., Boyle et al. (2015), Ebetürk (2021)

Notes: The table summarizes economic theories on the historical expansion of women's rights (largely based on the U.S. experience) and in the cross-section of countries, separately for each of the four types of rights. The question marks indicate gaps in the literature.

an undertaking, we now turn to empirical observations that can help inform such an analysis.

3.4 Women's Rights across Countries: Data Description and Regression Design

Given that much of the historical expansion of women's rights in the United States aligns with economic forces, we explore whether such forces similarly help explain why women's rights widely differ across countries today. We saw in Figure 3.1 and Table 3.1 above that women are still lacking many legal rights precisely in the poorest countries. Despite clear parallels to the historical development in high-income countries, there are also reasons to believe that additional factors—such as culture or international pressure—might be important for understanding contemporary heterogeneity in women's rights around the world. Outside of economics, scholars have stressed the significance of culture and specifically religion (Bayes and Tohidi 2001; Htun and Weldon 2011; Boyle et al. 2015; Ebetürk 2021). Another possibility is that growing enlightenment underlies the expansion of women's rights (Hyland et al., 2020). If so, women's rights should arguably have spread largely in parallel in many countries. We explore these determinants empirically, separately for each of the four types of women's rights.

3.4.1   Four Indices of Women's Rights

We consider four types of legal rights for women: economic, political, labor market, and body rights. We primarily use the World Bank's Women, Business, and the Law (WBL) database to create a rights index for each of these four areas, but augment it with information from other sources. The WBL data includes annual information on 35 legal rights for a panel of 190 countries[14] for the time period of 1970 to 2021 and specifies for each year whether a specific legal right existed for women in each country (see Hyland et al. 2020 for a comprehensive description of the database). We do not use information on all 35 legal rights, but rather select those that relate to our four areas. Appendix Table 3.A.1 provides a complete list of the rights included in each of the four indices and their data sources. Since we code the legal rights as zero or one (either women have or do not have that right in a given country at a certain point in time), the index can be interpreted as the fraction of legal rights that women have in that area. We scale each index to lie between 0 and 100.

Both our economic and labor rights indices rely exclusively on information from the WBL database. To construct the economic rights index, we use 16 questions related to mobility, marriage, entrepreneurship, and assets. Our labor rights index instead relies on six different legal rights related to equality in the workplace and pay, such as the right to work the night shift in the same way as men or whether discrimination in employment based on gender is legally prohibited.[15]

As the WBL data does not include information on laws related to women's political rights, we use information from other sources. Specifically, our political rights index is based on whether women have the right to vote, the right to be elected, and whether there is a sizeable representation of women in the national parliament. Admittedly, this last factor is not exactly a political right, though it does help capture the extent to which women are involved in legislation. Information on the right to vote and the right to be elected is based on Skaaning et al. (2015). Meanwhile, that on women in parliament comes from the cross-national data set compiled by Paxton et al. (2008), who record the year in which a country first attained a 20% or greater proportion of women in their parliament. A country may drop below this threshold in later years. We construct an indicator equal to zero as long as the share of women remains below 20%, which then takes a value of one in the first year in which women make up 20% or more, remaining one thereafter. Since this data is only available until 2003, we extend the indicator using information from the World Bank's World Development Indicators on the proportion of

---

[14] Throughout the sample period, some federal unions (e.g., the Soviet Union) dissolved and, as a result, new countries entered the sample. In other cases, two countries merged (e.g., East and West Germany). Appendix 3.A.3 summarizes how the WBL database deals with such cases.

[15] We exclude legislation related to sexual harassment in the workplace and include it instead in our body rights index.

seats held by women in national parliaments (World Bank, 2022b).

Finally, our body rights index combines data from the WBL and Boyle et al. (2015) to include six legal rights specifically related to a woman's own body.[16] From the WBL database, we use information on whether there is legislation against sexual harassment in employment, whether legal or civic remedies against sexual harassment exist, and whether legislation addressing domestic violence is in place. We then add information on abortion rights, based on Boyle et al., who collect data on seven different grounds for which an abortion might be legal.[17] We define two broad legal rights related to abortion: whether abortion is legal to save a mother's life or physical health or legal for any other reason, including simply a woman's request. These measures allow us to distinguish between a more narrow and a broader definition of abortion rights. As the data from Boyle et al. is available only until 2009, we supplement it with information from the World Population Policies Database in 2011, 2013, and 2015, and the UN Population Division in 2017 (United Nations, 2022a,b). Finally, we use contraceptive prevalence as a proxy for the right to reproductive autonomy.[18] Specifically, we employ data from the World Development Indicators database on the fraction of 15-49-year-old married women who use any method of contraception (World Bank, 2022b). Based on this information, we define an indicator that is equal to one if this fraction is equal to or above 45%, which is the average fraction of women using any method of contraception across all countries and years in our data.

Figure 3.3 plots the four women's rights indices over time as an average across all countries.[19] Notably, women's rights substantially increased in all four areas over the sample period.[20] Yet, there are clear differences between the different types of rights. Already in 1970, economic rights were well established in many countries, as were political rights with an index close to 60 in 1970. Both political and economic rights continued to increase, with the rise in political rights being more pronounced and notably surpassing the economic index in 2010.

---

[16] A more extensive analysis might include additional rights, such as laws against rape and specifically marital rape, regulations prohibiting female genital cutting, and whether or not child marriage is banned. We do not include these other dimensions in our analysis due to data limitations. Though some data does exist in this regard, it is typically available only a subset of countries and selected years, which would have severely impacted the number of observations in our analysis.

[17] These include to save a mother's life, protect a mother's physical health, protect a mother's mental health, pregnancy as a result of rape, fetal impairment, socioeconomic hardship, and a woman's request to have an abortion.

[18] Certain types of contraception, such as the pill, have historically been illegal in many countries.

[19] Since our indices contain different numbers of laws, a one point change does not refer to equal changes in the number of rights implemented across indices. For instance, an increase of 6.25 points in the economic rights index can be interpreted as, on average, one more economic law aiming at gender equality across countries. To reach the same change in terms of the number of laws, labor and body rights would need to increase by 16.67 points and political rights by 33.33 points. Nevertheless, changes in the index can be compared across areas as they always measure the change in the average percentage of rights in the respective index.

[20] Hyland et al. (2020) likewise document substantial progress in women's labor-related rights over the same period.

Figure 3.3: Worldwide Increase in Four Women's Rights Indices
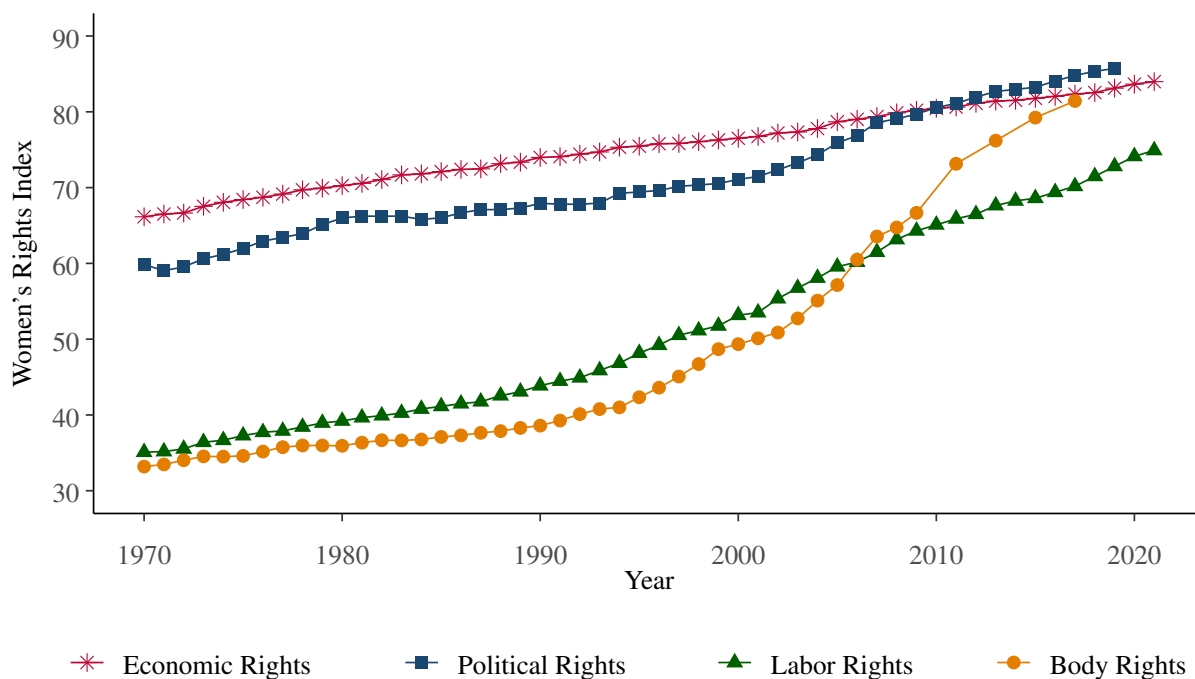
Notes: We construct four indices to measure economic, political, labor, and body rights using the WBL Database (World Bank, 2022a) as our main source, supplemented with information from various other sources. The construction of the indices is described in the text and the data sources are summarized in Appendix Table A.1.

In contrast, rights related to the labor market and the female body were much less common in 1970: the body rights index is 33 and the labor rights index 35. Labor rights gradually increased over the entire time period, reaching close to 75 points by 2021. Body rights were rather stagnant throughout the 1970s and 1980s and then started to sharply rise in the late 1990s. In 2006, the body rights index surpassed the labor rights index. The ordering of the indices throughout most of the considered time period is consistent with the historic sequence in the expansion of women's rights in the United States: economic rights first, followed shortly thereafter by political rights, with rights related to the labor market coming much later, and body rights only relatively recently.

We also construct an overall index comprised of the 25 rights from the WBL data, plus female suffrage and contraceptive prevalence. We exclude data on abortion rights, the right to be elected, and women in parliament to reduce the number of missing observations. Figure 3.4 displays how this overall index differs across countries at three points in time, revealing considerable geographic differences. In 1971, women in Africa, Latin America, the Middle East, and parts of Europe (most notably Spain) had few rights, while women in North America, Australia, and the northern parts of Europe already had many rights. By 1991, women's rights had improved in many places, though little progress had been made in most of Africa and the Middle

East. By 2017, the legal position of women had most strongly improved in South America, with North Africa and the Middle East being slowest to adopt different women's rights.[21]

3.4.2 Regression Design

To explore the drivers of women's rights, we combine our legal rights data with information on economic variables, namely GDP per capita, the total fertility rate (TFR), and female labor force participation (FLFP). We choose these variables because they relate to the different economic channels for expanding women's rights discussed above. The parental altruism channel links women's rights to family changes and specifically education investments in children, which is closely related to the fertility rate through the quantity-quality tradeoff.[22] The income channel concerns the implications of women's rights for women's labor supply, providing a link with the FLFP rate.

We do not include a separate variable for the public policy channel, as changes in political preferences are driven by the same changes to the role of families and women's labor supply that underlie the parental altruism and the income channels. Given that the public policy channel applies specifically to political rights, our economic variables can therefore be interpreted as also capturing the public policy channel in our results on political rights. Similarly, we do not include a separate variable for the bargaining channel. Apart from being difficult to measure, this channel relating to basic distributional conflict between women and men is arguably always present and less drastically transformed over time compared to the parental altruism and income channels.[23] This leads us to focus on the changing role of the other channels as drivers of political reform.

In addition to the relevance of economic forces, we explore two other determinants of women's rights emphasized in the literature: culture and enlightenment. We consider the influence of culture in two ways. Since culture can be thought of as invariant over shorter time horizons, we include country fixed effects to control for culture and other country-specific factors in some of our regressions. With regard to more specific cultural influences, Boyle et al. (2015) and Ebetürk (2021) point out that religion is particularly important for the expansion of women's rights (or the lack thereof). Hence, we include dummy variables for the religious majority in a given country and year in some specifications.[24] Table A.2 in the Appendix lists

---

[21] In their analysis of labor rights specifically, Hyland et al. (2020) similarly document an uneven progress in women's rights across regions.

[22] In a robustness exercise, we also link women's rights directly to children's education by using data on secondary school enrollment, see Section 3.5.6.

[23] Of course, the relative bargaining power in marriage has likely changed considerably over time due to, for example, changes in FLFP. However, such changes would not alter the preference of men for women to have little bargaining power, which is what our bargaining power channel is about.

[24] Online Appendix Table B.1 provides an overview of religious majorities in the year 2000 for each of the coun-

Figure 3.4: The Development of Women's Rights Across the World

(a) 1971

(b) 1991

(c) 2017
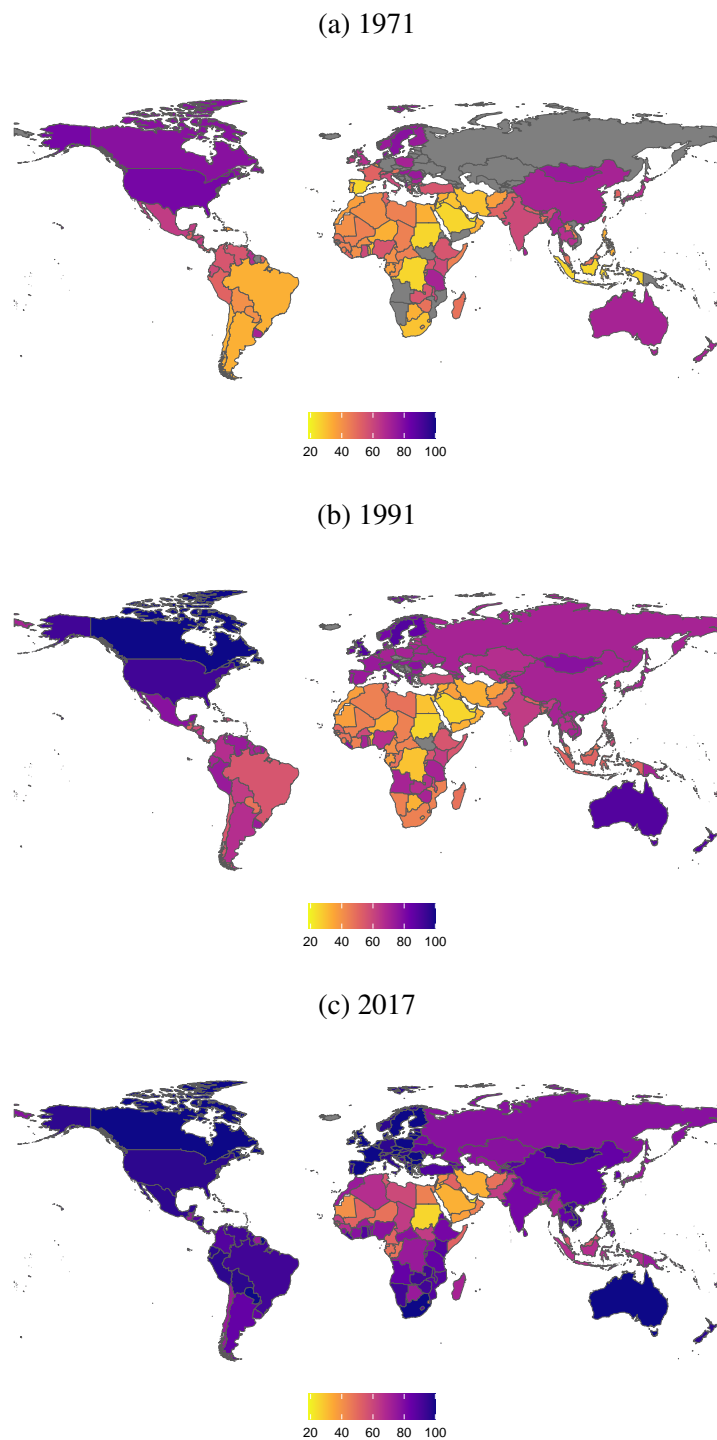
Notes: We construct an overall index that includes all the different individual rights comprising the economic, political, labor market, and body rights indices (see also Appendix Table A.1). The overall index expresses the percentage of laws that exist in a given country year and can range from 0 to 100.

all explanatory variables, along with their sources.

If culture or religion were the primary reasons for why women's rights differ across countries, one would expect the dummy variables for religious majorities or country fixed effects to account for much of the variation in the data. In regressions with religion effects, we include dummy variables for Muslim, Catholic, and Buddhist majorities. The omitted category consists of all other religions and non-religious majorities. To assess how economic development interacts with culture, in some regressions we also include interaction terms between the religious majority dummies and GDP per capita.

Examining the enlightenment channel, that is, the idea that the expansion of women's rights is a part of a general trend characterized by growing recognition of the rights of previously excluded groups, which progresses independently of economic changes, poses more of a challenge. As there are no direct empirical measures of such a global movement, we assess the role of the enlightenment channel by including time fixed effects in some of our regressions. If the spread of women's rights is due to a general increase in enlightened views in society, this should be reflected in a rise in these fixed effects over time. If growing enlightenment, independent of local economic conditions, was the primary force underlying the expansion of women's rights, we would expect time fixed effects to absorb most of the variation in the data.

Depending on data availability, our regressions cover slightly different time periods. For instance, whereas our economic and labor rights indices are available from 1970 to 2021, the political and body rights indices end in 2019 and 2017, respectively. Furthermore, information on Muslim and Buddhist population shares is only available until 2013.

3.5 Women's Rights across Countries: Empirical Results

We now turn to the findings from our empirical analysis, discussing the role of economic and cultural factors for the development of women's economic, political, labor, and body rights separately. We conclude this section with several robustness checks as well as an assessment of additional potentially relevant drivers of these rights. Tables 3.3–3.6 summarize the results, and we provide more detailed versions in Appendix 3.A.4.

3.5.1 Economic Rights

Table 3.3 reports the results for the economic rights index. Column (1) shows that there is a positive correlation between economic development (as measured by per capita GDP) and women's economic rights. However, the correlation is fairly small, and the low R^2 suggests

---

tries included in our analysis.

that little of the variation in women's economic rights across countries can be explained by economic development per se. Next, we include the TFR and FLFP as additional explanatory variables in column (2). The R^2 suggests that half of the variation in the data can be accounted for by these three economic variables. The effects are sizeable. The estimated coefficients from column (2) imply that a decline in the TFR of one child is related to an increase in the economic rights index of 7.15, which roughly corresponds to the difference in women's economic rights between India and the United States in 2000. Similarly, an increase in FLFP of 10 percentage points is associated with a 5.3 points higher economic index, which roughly corresponds to the worldwide increase in women's economic rights between 1970 and 1980. Both findings point to the income channel and the parental altruism channel as important mechanisms for explaining the variation in women's economic rights across countries.

Table 3.3: Economic Rights – Summary of Regression Results

	Dependent Variable: Economic Rights Index				
	(1)	(2)	(3)	(4)	(5)
A. Economic Variables					
GDP p.c. (in 1000s)	0.32***	−0.16**	0.04	−0.17**	−0.05
Total Fertility Rate		−7.15***	−1.71***	−6.75***	−6.53***
Female LFP (15-64)		0.53***	0.14***	0.45***	0.43***
B. Interactions with Religion					
Majority Catholic × GDP p.c.					0.12
Majority Muslim × GDP p.c.					−0.75***
Majority Buddhist × GDP p.c.					−0.26
Time Fixed Effects	No	No	Yes	Yes	Yes
Country Fixed Effects	No	No	Yes	No	No
Religion Fixed Effects	No	No	No	Yes	Yes
Observations	7,796	5,428	5,428	4,045	4,045
Adjusted R^2	0.057	0.513	0.919	0.541	0.583

Notes: All regressions include a dummy variable that is equal to one for the year in which a country newly gained independence. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. For a full set of results including standard errors, refer to Appendix Table A.3.

Columns (3) and (4) measure the effect of culture. We first include country and time fixed effects. While these explain a substantial share of variation in the data, the effects of the TFR and FLFP remain significant.[25] To compare the role of economic versus cultural factors, we assess the explanatory power of religion without including any economic variables (Appendix Table A.3, column (6)). Religion alone explains roughly 30% of the variation in economic rights across countries and time. If we include both religion and economic factors, the adjusted R^2

---

[25] Results are similar without time fixed effects, as we show in the Appendix.

increases to around 55%. In contrast, religion adds less than 5 percentage points of explanatory power when compared to a specification in which we control for economic factors (column (2) versus column (4)). Thus, the effect of religion mainly operates through differences in FLFP and TFR across religions.

Finally, we add interaction terms between the dummy variables for religious majorities and GDP per capita. The coefficients on TFR and FLFP barely change, but we do find significant negative coefficients for the interaction of Islam and Buddhism (but not Catholicism) with GDP per capita. Thus, even though religion by itself does not appear to explain much, it interacts with economic development, such that certain religions slow down the expansion of women's rights through economic channels.

To look at the role of enlightenment, we plot the estimated time fixed effects of the specification in column (5) in Figure A.1 in the Appendix. The graphical illustration shows that there are no significant time fixed effects for economic rights after 1970. This suggests that, if the enlightenment channel has played any role at all in the development of economic rights, it must have been prior to 1970.

Overall, at least at the level of correlations, the cross-country results for economic rights indicate a substantial role of two economic channels, namely the parental altruism channel and the income channel. Religion appears to be relatively unimportant, except for the fact that certain religions lower the positive effect of economic development on women's rights. There is little evidence of a general enlightenment trend driving women's economic rights in the time period considered.

3.5.2   Political Rights

We repeat the analysis for the political rights index in Table 3.4. While political rights are not correlated with economic development as proxied by GDP per capita (column (1)), we do observe significant correlation with TFR and FLFP (column (2)). That said, the variation in the data accounted for by economic variables is lower compared to economic rights, as reflected by the lower adjusted R^2. The estimated coefficients are also smaller. A fertility decline of one child per woman is associated with a 2.45 points higher political rights index, which is half in size compared to the association with the economic rights index. The relation between FLFP and political rights is similarly lower compared to economic rights. Adding country and time fixed effects instead explains much of the variation in the data (column (3)), while including religion rather than country fixed effects in column (4) gives almost the same results as column (2). So again, while culture seems to be important, there appear to be dimensions of culture (or other country-specific factors) not captured by religion or growing enlightenment. When we include the interaction of religion and economic development in column (5), the adjusted

R^2 increases substantially to about 0.3. As for economic rights, the coefficients suggest that in countries with Muslim or Buddhist majorities the positive effect of economic development on women's rights is diminished. This interaction between economic development and religion is quantitatively more important for political compared to economic rights. A plot of the estimated time dummies of column (5) suggests that a general enlightened trend towards more political rights may have played a role only from the early 2000s onward (see Figure A.1 in the Appendix).

Table 3.4: Political Rights – Summary of Regression Results

	Dependent Variable: Political Rights Index				
	(1)	(2)	(3)	(4)	(5)
A. Economic Variables					
GDP p.c. (in 1000s)	0.11	−0.06	−0.06	−0.07	0.17*
Total Fertility Rate		−2.45***	−1.36***	−1.75**	−1.51**
Female LFP (15-64)		0.34***	0.35***	0.30***	0.26***
B. Interactions with Religion					
Majority Catholic × GDP p.c.					−0.22
Majority Muslim × GDP p.c.					−1.24***
Majority Buddhist × GDP p.c.					−0.42**
Time Fixed Effects	No	No	Yes	Yes	Yes
Country Fixed Effects	No	No	Yes	No	No
Religion Fixed Effects	No	No	No	Yes	Yes
Observations	7,138	5,167	5,167	3,981	3,981
Adjusted R^2	0.009	0.140	0.757	0.181	0.305

Notes: All regressions include a dummy variable that is equal to one for the year in which a country newly gained independence. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. For a full set of results including standard errors, see Appendix Table A.4.

3.5.3   Labor Rights

The results for the impact of economic variables and religion on labor market rights are largely similar to those for economic rights. Table 3.5 shows that there is a positive correlation between economic development and labor market rights for women. However, in contrast to economic and political rights, per capita GDP continues to be significant even after including TFR and FLFP. The estimated coefficients in column (2) imply that a reduction in the TFR of one child is associated with a 4.78 point increase in the labor index, which is larger than for the political rights index but somewhat smaller compared to the economic rights index. The estimated coefficient on FLFP is 0.52, which is comparable in magnitude to the effect on the economic rights index. As before, Muslim and Buddhist religious majorities, but not Catholic majorities, reduce

the positive effect of economic development on women's labor market rights. Finally, the estimated time fixed effects are positive, especially from the mid-1990s onward (see Figure A.1 in the Appendix), pointing to a possible role of enlightenment for the expansion of women's labor rights.

Table 3.5: Labor Rights – Summary of Regression Results

	Dependent Variable: Labor Rights Index				
	(1)	(2)	(3)	(4)	(5)
A. Economic Variables					
GDP p.c. (in 1000s)	0.62***	0.26**	−0.26***	0.35***	0.60***
Total Fertility Rate		−4.78***	2.24***	−2.59***	−2.37***
Female LFP (15-64)		0.52***	0.06	0.31***	0.28***
B. Interactions with Religion					
Majority Catholic × GDP p.c.					−0.27
Majority Muslim × GDP p.c.					−1.23***
Majority Buddhist × GDP p.c.					−0.69***
Time Fixed Effects	No	No	Yes	Yes	Yes
Country Fixed Effects	No	No	Yes	No	No
Religion Fixed Effects	No	No	No	Yes	Yes
Observations	7,796	5,428	5,428	4,045	4,045
Adjusted R^2	0.123	0.299	0.826	0.380	0.441

Notes: All regressions include a dummy variable that is equal to one for the year in which a country newly gained independence. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. For a full set of results including standard errors, see Appendix Table A.5.

3.5.4 Rights over Own Body

Finally, Table 3.6 repeats the analysis for body rights. As before, economic development, fertility, and FLFP are strongly linked to increasing female body rights. The estimated coefficient on the TFR in column (2) is -5.82, meaning that the total fertility rate and body rights are linked more strongly than in the case of the political and labor rights indices, but that the link is weaker compared to the economic rights index. The association with FLFP is lower for the body rights index than for any of the other rights. Taken together, economic variables account for less of the variation compared to economic and labor rights, but more relative to political rights.

An interesting difference is that growing enlightenment (assessed through time fixed effects) and religious majorities seem to play a larger role than for any of the other rights. Indeed, when comparing columns (2) and (4), we see that the adjusted R^2 increases by 60%. This effect is mainly driven by time fixed effects. Appendix Figure A.1 plots the estimated time fixed effects, which are essentially zero until 1990 and then rise sharply thereafter, closely aligning with the

worldwide acceleration of body rights in the mid-1990s. Thus, a rising worldwide awareness for individual rights may indeed be an important driver of women's body rights specifically.

Table 3.6: Rights over Own Body – Summary of Regression Results

	Dependent Variable: Body Rights Index				
	(1)	(2)	(3)	(4)	(5)
A. Economic Variables					
GDP p.c. (in 1000s)	0.57***	0.14*	0.27***	0.22**	0.33***
Total Fertility Rate		−5.82***	2.66***	−4.30***	−4.06***
Female LFP (15-64)		0.32***	0.10**	0.20***	0.18**
B. Interactions with Religion					
Majority Catholic × GDP p.c.					0.07
Majority Muslim × GDP p.c.					−0.52***
Majority Buddhist × GDP p.c.					−0.18
Time Fixed Effects	No	No	Yes	Yes	Yes
Country Fixed Effects	No	No	Yes	No	No
Religion Fixed Effects	No	No	No	Yes	Yes
Observations	5,578	3,950	3,950	3,492	3,492
Adjusted R^2	0.106	0.268	0.776	0.430	0.445

Notes: All regressions include a dummy variable that is equal to one for the year in which a country newly gained independence. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. For a full set of results including standard errors, see Appendix Table A.6.

The relatively larger role of religion in these regressions is consistent with empirical studies in sociology. Boyle et al. (2015), for example, show that religion is strongly correlated with the level of abortion rights in a country. Interestingly, we do not find a significant effect of Catholic majorities on female body rights. This may be due to the fact that our body rights index includes several body rights dimensions, while Boyle et al. analyze abortion rights specifically.[26] Abortion laws are unique in the sense that they legislate the tradeoff between women's own health, including their reproductive autonomy, versus the life or health of an unborn child. Abortion laws must accordingly take a legal stance on when life itself starts. Catholicism clearly identifies the moment of conception as the relevant reference point. It is thus no surprise that countries with Catholic majorities are negatively associated with abortion rights, but not necessarily with a more extended notion of body rights that includes domestic violence or sexual harassment. Finally, we find that Muslim, but not Buddhist, majorities also diminish the positive effect of economic development on body rights.

---

[26] Boyle et al. (2015) distinguish between three different types of abortion rights and find that Catholicism plays a particularly important role in blocking abortion due to fetal impairment or mother's mental health but is less relevant for the right to an abortion due to rape.

3.5.5   Summary and Interpretation of Empirical Findings

The empirical analysis suggests that women's rights are strongly associated with economic development. Only political rights are not correlated with GDP per capita, although they are associated with other economic variables. In terms of magnitudes, the economic channels explain the largest share of the variation in the data for the case of economic rights (with an R^2 of 0.51), followed by labor rights (0.30), then body rights (0.27), with political rights last (0.14).

Two key findings emerge. First, the FLFP rate is positively associated with all four types of women's rights, suggesting that the income channel could be an important driver for the expansion of women's rights. This channel predicts that political preferences for women's rights depend not only on one's own labor supply, but also that of one's spouse. As a result, political support for women's rights is not simply a function of gender. Men with labor market skills that are complementary to women's benefit from an expansion in women's rights, as do men whose spouses would increase their earnings when rights are extended. Not surprisingly, the income channel seems most important for economic and labor rights, with estimated coefficients around 0.5, compared to 0.3 for the political and body rights (referring to column (2) of Tables 3.3–3.6 in each case).

Second, the TFR is also significantly negatively related to all types of women's rights, implying that parental altruism may be another important channel underlying the expansion of women's rights. If parents care about the human capital of their children and women's rights affect investments in children, parents of either gender may be in favor of extending women's rights to improve child outcomes and child welfare. The strength of this channel depends on the return to human capital, where a high return is also associated with low fertility through the quantity-quality tradeoff. Again, this channel seems most important for economic rights, with an estimated coefficient of around -7 in column (2) of Table 3.3, closely followed by body rights and labor rights with coefficients of about -6 and -5, respectively. The quantitative effect is smallest for political rights, with a coefficient of roughly -2.5. The large impact of changes in fertility on body rights is perhaps unsurprising, given that female body rights such as abortion and rape laws are naturally linked to childbearing.

Even though our analysis suggests that economic forces are key in explaining cross-country differences in women's rights, we also find that religion interacts with economic development and can diminish the effect of economic channels that increase women's rights. Specifically, we find that Muslim majorities are associated with fewer women's rights, independently of the type of right. The interaction between economic development and religion is particularly important for body rights. Time fixed effects are similarly most relevant for body rights, suggesting that general enlightenment may have played a role in the expansion of own body rights in recent decades.

3.5.6 Robustness Checks

We run several alternative specifications of our main regressions to test the robustness of our results. Tables B.2 – B.5 in the Online Appendix display these additional findings for the economic, political, labor, and body rights indices, with columns (2) to (6) comparing the results to our main specification, given in column (1).[27]

In our main regressions, we use the maximum number of observations available for each specification, and as such use different samples across regressions and indices. Our results are robust to using a consistent sample and not driven by differences in data availability, as shown in column (2).

We also control for OPEC membership to capture the fact that some countries have become wealthy without having experienced a more typical development process of structural change and increasing demand for human capital. While OPEC membership is significantly negatively associated with economic rights, it is not significantly correlated with other rights, nor does it affect the sign and significance of our main coefficients, see column (3).

Similarly, we look at a specification without time fixed effects (but with religion) to assess the extent to which time-specific effects interact with the effects of economic and cultural characteristics on the respective rights. Comparing the two specifications in columns (1) and (4), we see that time fixed effects do not substantially alter the sign and magnitude of coefficients, suggesting that any potential time or enlightenment effects are uncorrelated with the effects of economic and cultural variables.

Finally, in column (5), we replace the TFR with secondary school enrollment as a proxy for the quantity-quality tradeoff. In line with our theory, the TFR and secondary school enrollment have opposite signs, and secondary school enrollment is highly significant across all indices. At the same time, the sign and significance of other contributing factors are largely consistent across specifications.

3.5.7 Factors Specific to Individual Laws

We now investigate the effects of several additional factors on women's economic, political, labor, and body rights by adding one additional explanatory variable at a time to our baseline specification with time and religion fixed effects and interaction terms, that is, to the last column in Tables 3.3–3.6.[28] Table 3.7 summarizes the sign of significant coefficients in these additional regressions for each of the four different indices. Detailed results are provided in the Online Appendix B.3.

---

[27] To be precise, column (1) always repeats column (5) of our main Tables 3.3–3.6.

[28] Sources for the additional explanatory variables are included in Table A.2 of the Appendix.

Table 3.7: Additional Factors Correlated with Women's Rights

Factor/Rights	Economic	Political	Labor	Own Body
Female Population Share				
Employment Share Agriculture		+		
Government Effectiveness				+
Membership in Int. Women NGOs		+		
Nb. of Conventions Ratified	+	+		+
International Pressure	+	+	+	
Women in Parliament $> 20\%$	+	n/a	+	

Notes: All regression results underlying this summary table can be found in Tables B.6 – B.9 of the Online Appendix.

Some scholars have argued that additional rights would be granted more willingly to women if they made up a relatively small share of the population, and thus such expansion would not affect the lives of men substantially (e.g., Jones 1991). That said, a larger share of women could be indicative of greater collective female bargaining power. In contrast to both theories, we find no significant association of the share of female citizens with any type of women's right.

Following Bertocchi (2011), we include the employment share in agriculture as an additional factor potentially associated with women's rights, the idea being that women have a comparative advantage in brain-based employment, such that their relative wages are higher in more industrialized settings where the role of intellectual work is more important than in agrarian societies. Lower shares of agricultural employment would be directly linked to a lower gender wage gap, and hence higher women's bargaining power, and thus possibly more rights. In our data, only political rights are significantly associated with employment in agriculture. However, against our prediction, this effect is positive.[29]

In her study on legislation against forced and child marriage, Ebetürk (2021) argues that legislation establishing women's rights is also likely related to state capacity. In order to recognize the absence of women's rights as well as to pass counteracting legislation, a state must have the capacity to identify the prevalence of inequalities between women and men. In line with this argument, Ebetürk finds a positive effect of the size of government on child labor laws. Relatedly, Poyker (2022) finds that female genital cutting is more persistent where governments are unstable. Using an index of government effectiveness by Teorell et al. (2013), we observe that this result holds for body rights more generally. However, we find no significant correlation of government effectiveness with economic, labor, or political rights.

Both Boyle et al. (2015) and Ebetürk (2021) emphasize that the transmission of global norms are key for the expansion of women's rights across countries. The degree to which they

---

[29] In her analysis of suffrage extensions in 22 countries during the period of 1870-1930, Bertocchi (2011) finds no significant effect of agricultural share on the right to vote and hypothesizes that this is due to the high correlation of per capita GDP and agricultural share.

affect women's rights in a given country may depend on two factors: (i) the number of linkages a country has to the rest of the world and (ii) the intensity of such linkages. Following Ebetürk, we measure this in several ways. First, we use the number of women's international NGOs in which residents of a given country are members. While Ebetürk finds no significant relationship between memberships in women's international NGOs and child marriage legislation, we document a positive effect for women's political rights (though not for any of the other rights). When using the number of ratified international conventions related to women and children, additional positive effects on women's economic and own body rights emerge. This is consistent with Ebetürk, who observes that a larger number of ratified conventions is associated with more legislation against child marriage.

Second, to measure the intensity of linkages to other countries, we hypothesize that neighboring countries are often culturally similar and economically interconnected. Hence, we would expect the influence of global norms in a neighboring country to play a larger role than those in a country that is geographically farther away. Such diffusion has been documented in other contexts, e.g., Spolaore and Wacziarg (2022) document cultural diffusion of fertility behavior, but this largely remains unexplored in the context of women's rights. We construct a variable that summarizes the average number of ratifications in the region to which a country geographically belongs. We call this measure "international pressure," as it captures the influence of global norms beyond the country itself. While our measure is positively associated with women's economic, political, and labor rights, there is no correlation with rights over one's own body. This may suggest that global norms are insufficient to affect deeply rooted norms and traditions related to women's body rights and that these rights are only granted once the country itself is ready to question traditional norms.

Finally, active female participation in the political process is likely to influence legislation related to women's rights. To test the effects of women's political representation, we look at the effects of the share of women in parliament. Specifically, we measure female representation using a dummy variable indicating whether the share of women in parliament has ever been above 20%. Table 3.7 shows that the presence of women in parliament is positively associated with both economic and labor rights, but does not affect women's rights over their own body.[30]

3.6   Conclusion and Directions for Future Work

Women's rights have progressively expanded around the world, yet sizeable gaps between the legal rights of women and men remain. In this paper, we investigate the political economy be-

---

[30] Since the women in parliament indicator is part of our political rights index, we cannot include it as an explanatory variable for political rights.

hind the expansion of four different types of women's rights: economic, political, labor market, and body-related rights. We argue that understanding who stands to gain or lose from such rights is key to elucidating the political process leading to the expansion of women's rights. To do so, we present a simple framework that illustrates four different mechanisms through which individual preferences about women's rights are affected: the bargaining power channel, the parental altruism channel, the income channel, and the public policy channel. We contend that changes in economic development (such as technological progress and structural change) can alter the tradeoffs implied by these four channels, resulting in a shift where the majority of people who initially voted against women's rights subsequently vote in favor. There is substantial evidence in the literature that economic forces have been important for the historical expansion of women's rights. Here, we present new empirical evidence suggesting that such forces may be equally critical in explaining cross-country differences in the legal position of women today. We find that our economic variables explain the largest share of cross-country variation for women's economic and labor rights and the least for their political rights. Economic channels also explain a sizeable share of the cross-country variation in women's body rights, with the impact of the fertility rate being particularly strong.

Many open questions remain. First and foremost, there are no economic theories focused on explaining today's differences in women's rights across countries. The existing economic models analyze the evolution of women's economic and political rights over time in a given country, mainly in the context of U.S. history. Second, little theoretical work assesses the different types of women's rights. For instance, in the United States, women's economic rights expanded about a century earlier than body rights, and our cross-country analysis suggests that the driving forces behind the expansion of these two rights may differ. Finally, many examples show that women's rights have not always progressed in a linear fashion. In some contexts, there have even occurred reversals of women's rights. For example, the United States Supreme Court granted a constitutional right to abortion for women in 1973 (Roe v Wade), but the ruling was overturned again in 2022, when the Supreme Court took a different stand on the interpretation of the U.S. Constitution (Dobbs v Jackson Women's Health Organization). Women's rights in Afghanistan have likewise oscillated back and forth over time, largely related to the Taliban regime. Do economic forces play a role in such reversals as well?

We abstract from the issue of enforcement in our analysis.[31] Yet, without robust enforcement the legal position of women for a given de jure law is much weaker. What explains differences in enforcement? How can we measure enforcement? These are questions left for future research. New data might help to further disentangle the causal relationships between laws, attitudes, and practice. The United Nations Development Program created the new Gender So-

---

[31] Acemoglu and Jackson (2017) offer a general theory of the interaction between social norms and the enforcement of laws; it would be interesting to apply their theory to women's rights specifically.

cial Norms Index, which documents gender norms along four dimensions (politics, education, economics/work, and physical integrity) across countries. Similarly, the OECD Gender, Institutions and Development Database distinguishes between law, attitudes, and practice on several dimensions of gender equality.

Finally, research documents that the expansion of women's rights in one dimension is often accompanied by a weakening of their positions in other dimensions, a phenomenon known as "male backlash." This captures the idea that when women advance in one area of life, they may pay for it by losing ground in another sphere. If so, more legal rights may not always universally improve the position of women, and there might be a tradeoff between formal legal rights for women and their well-being. While substantial evidence shows that more resources in the form of income transfers or encouraging women to work can lead to backlash, there has been little exploration of this idea in the context of women's legal rights. Studies focusing on resources more generally include Angelucci (2008), who observes in data from Mexico that large income transfers to women led to more domestic violence. Anderberg and Rainer (2013) develop a theory explaining why an increase in female wages can lead to more domestic violence where the husband tries to sabotage her work efforts. Similarly, Eswaran and Malhotra (2011) find that greater female autonomy increases domestic violence in India and contend that men use violence to increase their bargaining power. Bloch and Rao (2002) show that dowry violence increases in the wife's family income and develop a theory to explain this pattern. Meanwhile, Angelucci and Heath (2020) find that being the main earner is positively correlated with domestic violence among Congolese women. Domestic violence is not the only form of male backlash that women whose bargaining power has increased may experience. For instance, women who work longer hours than their husbands tend to compensate by doing more housework, and accordingly consuming less leisure and suffering in terms of life satisfaction (see Flèche et al. 2018 and Flèche et al. 2020). Wheaton (2021) is one of the few papers that analyze male backlash in the context of women's legal rights. The author finds that men reacted sharply to Equal Rights Amendments established in many U.S. states throughout the 1970s, expressing more negative attitudes toward gender equality immediately after their introduction. Sanders and Jenkins (2022) discusses more recent backlash to women's rights on the part of populist leaders. Considering the possibility of backlash is another promising direction for future research on the economics of women's rights.

Bibliography

Abrams, B. A. and Settle, R. F. (1999). Women's suffrage and the growth of the welfare state. *Public Choice*, 100(3–4):289–300.

Acemoglu, D. and Jackson, M. O. (2017). Social norms and the enforcement of laws. *Journal of the European Economic Association*, 15(2):245–295.

Acemoglu, D. and Robinson, J. A. (2000). Why did the West extend the franchise? Democracy, inequality and growth in historical perspective. *Quarterly Journal of Economics*, 115(4):1167–99.

Aidt, T. S. and Dallal, B. (2008). Female voting power: The contribution of women's suffrage to the growth of social spending in Western Europe (1869–1960). *Public Choice*, 134(3):391–417.

Aizer, A. (2010). The gender wage gap and domestic violence. *American Economic Review*, 100(4):1847–59.

Alesina, A., Giuliano, P., and Nunn, N. (2013). On the origins of gender roles: Women and the plough. *The Quarterly Journal of Economics*, 128(2):469–530.

Alshaikhmubarak, H., Geddes, R. R., and Grossbard, S. A. (2019). Single motherhood and the abolition of coverture in the United States. *Journal of Empirical Legal Studies*, 16(1):94–118.

Anderberg, D. and Rainer, H. (2013). Economic abuse: A theory of intrahousehold sabotage. *Journal of Public Economics*, 97:282–295.

Angelucci, M. (2008). Love on the rocks: Domestic violence and alcohol abuse in rural Mexico. *B.E. Journal of Economic Analysis and Policy*, 8(1):1–43.

Angelucci, M. and Heath, R. (2020). Women empowerment programs and intimate partner violence. *AEA Papers and Proceedings*, 110:610–614.

Bailey, M. J. (2006). More power to the pill: The impact of contraceptive freedom on women's life cycle labor supply. *The Quarterly Journal of Economics*, 121(1):289–320.

Bailey, M. J., Helgerman, T., and Stuart, B. A. (2021). How the 1963 Equal Pay Act and 1964 Civil Rights Act shaped the U.S. gender gap. Presented at NBER Summer Institute: Development of the American Economy, 15 July 2021.

Baker, C. N. (2022). The history of abortion law in the United States. *Our Bodies Ourselves Today Project*, Suffolk University. https://www.ourbodiesourselves.org/health-info/u-s-abortion-history/.

Barro, R. J. (1999). Determinants of democracy. *Journal of Political Economy*, 107(S6):S158–S183.

Bayes, J. and Tohidi, N. (2001). *Globalization, Religion and Gender: The Politics of Women's Rights in Catholic and Muslim Contexts*. Palgrave MacMillian.

Becker, A. (2019). On the economic origins of restrictions on women's sexuality. CESifo Working Paper No. 7770.

Bertocchi, G. (2011). The enfranchisement of women and the welfare state. *European Economic Review*, 55(4):535 – 553.

Blau, F. D. and Kahn, L. M. (2017). The gender wage gap: Extent, trends, and explanations. *Journal of Economic Literature*, 55(3):789–865.

Bloch, F. and Rao, V. (2002). Terror as a bargaining instrument: A case study of dowry violence in rural India. *American Economic Review*, 92(4):1029–1043.

Boserup, E. (1970). *Woman's Role in Economic Development*. George Allen and Unwin Ltd., London.

Bossen, L., Xurui, W., Brown, M. J., and Gates, H. (2011). Feet and fabrication: Footbinding and early twentieth-century rural women's labor in Shaanxi. *Modern China*, 37(4):347–383.

Boyle, E. H., Kim, M., and Longhofer, W. (2015). Abortion liberalization in world society, 1960–2009. *American Journal of Sociology*, 121(3):882–913.

Braun, S. and Kvasnicka, M. (2013). Men, women, and the ballot: Gender imbalances and suffrage extensions in the United States. *Explorations in Economic History*, 50(3):405–426.

Brooks, N., Kim, M., Boyle, E. H., and Longhofer, W. (2022). A post-Roe world? Why abortion battles in America won't halt reform abroad. *Foreign Affairs*, June 16.

Bursztyn, L., González, A. L., and Yanagizawa-Drott, D. (2020). Misperceived social norms: Women working outside the home in Saudi Arabia. *American Economic Review*, 110(10):2997–3029.

Chattopadhyay, R. and Duflo, E. (2004). Women as policy makers: Evidence from a randomized policy experiment in India. *Econometrica*, 72(5):1409–1443.

Cheung, S. N. S. (1972). The enforcement of property rights in children, and the marriage contract. *The Economic Journal*, 82(326):641–657.

Doepke, M., Foerster, H., Hannusch, A., and Tertilt, M. (2021). The political economy of female labor laws. Presented at the Annual Meetings of the Society for Economic Dynamics in Barcelona, July 2021.

Doepke, M. and Tertilt, M. (2009). Women's liberation: What's in it for men? *Quarterly Journal of Economics*, 124(4):1541–1591.

Doepke, M. and Tertilt, M. (2016). Families in macroeconomics. In *Handbook of Macroeconomics, Vol. 2*, chapter 23. North Holland.

Doepke, M. and Tertilt, M. (2019). Does female empowerment promote economic development? *Journal of Economic Growth*, 24(4):309–343.

Doepke, M., Tertilt, M., and Voena, A. (2012). The economics and politics of women's rights. *Annual Review of Economics*, 4(1):339–372.

Duflo, E. (2012). Women empowerment and economic development. *Journal of Economic Literature*, 50(4):1051–1079.

Ebetürk, I. (2021). Global diffusion of laws: The case of minimum age of marriage legislation, 1965–2015. *European Journal of Cultural and Political Sociology*, 8(3):294–328.

Edlund, L. and Pande, R. (2002). Why have women become left-wing? The political gender gap and the decline in marriage. *Quarterly Journal of Economics*, 117(3):917–961.

Eswaran, M. and Malhotra, N. (2011). Domestic violence and women's autonomy: Evidence from India. *Canadian Journal of Economics*, 44(4):1222–63.

Fernández, R. (2014). Women's rights and development. *Journal of Economic Growth*, 19(1):37–80.

Fernández-Villaverde, J., Greenwood, J., and Guner, N. (2014). From shame to game in one hundred years: An economic model of the rise in premarital sex and its de-stigmatization. *Journal of the European Economic Association*, 12(1):25–61.

Flèche, S., Lepinteur, A., and Powdthavee, N. (2018). Gender norms and relative working hours: Why do women suffer more than men from working longer hours than their partners? *AEA Papers and Proceedings*, 108:163–168.

Flèche, S., Lepinteur, A., and Powdthavee, N. (2020). Gender norms, fairness and relative working hours within households. *Labour Economics*, 65:101866.

Foerster, H. (2021). Untying the knot: How child support and alimony affect couples' decisions and welfare. CRC Discussion Paper No. 115, University of Mannheim, April 2020.

Frank, D. J., Hardinge, T., and Wosick-Correa, K. (2009). The global dimensions of rape-law reform: A cross-national study of policy outcomes. *American Sociological Review*, 74(2):272–290.

Funk, P. and Gathmann, C. (2014). Gender gaps in policy making: Evidence from direct democracy in Switzerland. *Economic Policy*, 30(81):141–181.

Geddes, R. and Lueck, D. (2002). The gains from self-ownership and the expansion of women's rights. *American Economic Review*, 92(4):1079–1092.

Geddes, R., Lueck, D., and Tennyson, S. (2012). Human capital accumulation and the expansion of women's economic rights. *The Journal of Law & Economics*, 55(4):839–867.

Geddes, R. R. and Tennyson, S. (2013). Passage of the married women's property acts and earnings acts in the United States: 1850 to 1920. *Research in Economic History*, 29:145–189.

Giuliano, P. (2022). Gender and culture. *Oxford Review of Economic Policy*, forthcoming.

Godefroy, R. (2019). How women's rights affect fertility: Evidence from Nigeria. *The Economic Journal*, 129(619):1247–1280.

Goldin, C. (1988a). Marriage bars: Discrimination against married women workers, 1920's to 1950's. NBER Working Paper No. w2747.

Goldin, C. (1988b). Maximum hours legislation and female employment: A reassessment. *Journal of Political Economy*, 96(1):189–205.

Goldin, C. and Katz, L. F. (2002). The power of the pill: Oral contraceptives and women's career and marriage decisions. *Journal of Political Economy*, 110(4):730–770.

Goldin, C., Katz, L. F., and Kuziemko, I. (2006). The homecoming of American college women: The reversal of the college gender gap. *Journal of Economic Perspectives*, 20(4):133–156.

Hazan, M., Weiss, D., and Zoabi, H. (2019). Women's liberation as a financial innovation. *The Journal of Finance*, 74(6):2915–2956.

Hazan, M., Weiss, D., and Zoabi, H. (2021). Women's liberation, household revolution. CEPR Discussion Paper 16838.

Hecker, E. A. (1971). *A Short History of Women's Rights*. Greenwood Press, Westport.

Htun, M. and Weldon, S. L. (2011). State power, religion, and women's rights: A comparative analysis of family law. *Indiana Journal of Global Legal Studies*, 18(1, Article 7):145–165.

Htun, M. and Weldon, S. L. (2018). *The Logics of Gender Justice: State Action on Women's Rights around the World*. Cambridge University Press, Cambridge.

Huber, J. (1976). Towards a sociotechnological theory of the women's movement. *Social Problems*, 23(4):371–88.

Hunt, J. C. and Rubin, P. H. (1980). The economics of the women's movement. *Public Choice*, 35(3):287–295.

Hyland, M., Djankov, S., and Goldberg, P. K. (2020). Gendered laws and women in the workforce. *American Economic Review: Insights*, 2(4):475–90.

Jayachandran, S. (2015). The roots of gender inequality in developing countries. *Annual Review of Economics*, 7(1):63–88.

Jones, E. B. (1991). The economics of women suffrage. *Journal of Legal Studies*, 20(2):423–37.

Khan, B. Z. (1996). Married women's property laws and female commercial activity: Evidence from United States patent records, 1790-1895. *The Journal of Economic History*, 56(2):356–388.

Landes, E. M. (1980). The effect of state maximum-hours laws on the employment of women in 1920. *Journal of Political Economy*, 88(3):476–494.

Lipset, S. M. (1959). Some social requisites of democracy: Economic development and political legitimacy. *The American Political Science Review*, 53(1):69–105.

Lise, J. and Yamada, K. (2019). Household sharing and commitment: Evidence from panel data on individual expenditures and time use. *Review of Economic Studies*, 86(5):2184–2219.

Lizzeri, A. and Persico, N. (2004). Why did the elites extend the suffrage? Democracy and the scope of government, with an application to Britain's "Age of Reform". *The Quarterly Journal of Economics*, 119(2):707–765.

Lott, J. R. and Kenny, L. W. (1999). Did women's suffrage change the size and scope of government? *Journal of Political Economy*, 107(6):1163–1198.

Manser, M. and Brown, M. (1980). Marriage and household decision-making: A bargaining analysis. *International Economic Review*, 21(1):31–44.

Marchingiglio, R. and Poyker, M. (2019). The employment effects of gender-specific minimum

wage. Working Paper 290, The University of Chicago Booth School of Business, George J. Stigler Center for the Study of the Economy and the State.

Mazzocco, M. (2007). Household intertemporal behaviour: A collective characterization and a test of commitment. *Review of Economic Studies*, 74(3):857–895.

McBride, D. E. and Parr, J. A. (2010). *Women's Rights in the USA: Policy Debates and Gender Roles*. Taylor & Francis Group.

Miller, G. (2008). Women's suffrage, political responsiveness, and child survival in American history. *Quarterly Journal of Economics*, 123(3):1287–1327.

Murtin, F. and Wacziarg, R. (2014). The democratic transition. *Journal of Economic Growth*, 19(2):141–181.

Myers, C. K. (2017). The power of abortion policy: Reexamining the effects of young women's access to reproductive control. *Journal of Political Economy*, 125(6):2178–2224.

Olivetti, C. and Petrongolo, B. (2016). The evolution of gender gaps in industrialized countries. *Annual Review of Economics*, 8:405–434.

Oswald, A. J. and Powdthavee, N. (2010). Daughters and left-wing voting. *Review of Economics and Statistics*, 92(2):213–227.

Pande, R. and Roy, H. (2021). "If you compete with us, we shan't marry you": The (Mary Paley and) Alfred Marshall lecture. *Journal of the European Economic Association*, 19(6):2992–3024.

Paxton, P., Green, J., and Hughes, M. M. (2008). Women in parliament, 1945-2003: Cross-national dataset. https://doi.org/10.3886/ICPSR24340.v1. Inter-university Consortium for Political and Social Research.

Poyker, M. (2022). Regime stability and the persistence of traditional practices. *The Review of Economics and Statistics*, forthcoming.

Salmon, M. (1986). *Women and the Law of Property in Early America*. The University of North Carolina Press, Chapel Hill.

Sanders, R. and Jenkins, L. D. (2022). Control, alt, delete: Patriarchal populist attacks on international women's rights. *Global Constitutionalism*, page 1–29.

Skaaning, S.-E., Gerring, J., and Bartusevicius, H. (2015). A lexical index of electoral democracy. *Comparative Political Studies*, 48(12):1491–1525.

Slotwinski, M. and Stutzer, A. (2022). Women leaving the playpen: The emancipating role of female suffrage. *The Economic Journal*, forthcoming.

Spolaore, E. and Wacziarg, R. (2022). Fertility and modernity. *The Economic Journal*, 132(642):796–833.

Stevenson, B. and Wolfers, J. (2006). Bargaining in the shadow of the law: Divorce laws and family distress. *The Quarterly Journal of Economics*, 121(1):267.

Svaleryd, H. (2009). Women's representation and public spending. *European Journal of Political Economy*, 25(2):186–198.

Teele, D. L. (2018). *Forging the Franchise: The Political Origins of the Women's Vote*. Princeton University Press.

Teorell, J., Sundström, A., Holmberg, S., Rothstein, B., Alvarado Pachon, N., and Dalli, C. M. (2013). The quality of government basic dataset. University of Gothenburg: The Quality of Government Institute. https://www.gu.se/en/quality-government.

Tertilt, M. (2006). Polygyny, women's rights, and development. *Journal of the European Economic Association*, 4(2-3):523–530.

United Nations (2022a). World population policies database (2011, 2013, 2015). https://esa.un.org/PopPolicy/wpp_datasets.aspx (accessed August 9, 2022).

United Nations (2022b). World population policies database (2017). https://www.un.org/development/desa/pd/data/world-population-policies (accessed August 9, 2022).

Voena, A. (2015). Yours, mine, and ours: Do divorce laws affect the intertemporal behavior of married couples? *American Economic Review*, 105(8):2295–2332.

Washington, E. L. (2008). Female socialization: How daughters affect their legislator fathers' voting on women's issues. *American Economic Review*, 98(1):311–332.

Wheaton, B. (2021). Essays on political economy and macroeconomics. Doctoral dissertation, Chapter 1, Harvard University Graduate School of Arts and Sciences.

World Bank (2022a). Women business and the law data for 1971–2022. https://wbl.worldbank.org/content/dam/sites/wbl/documents/2021/02/WBL1971-2022%20Dataset.dta (accessed June 7, 2022).

World Bank (2022b). World development indicators. https://databank.worldbank.org/source/world-development-indicators (accessed July–August 2022).

Yalom, M. (2001). *A History of the Wife*. HarperCollins Publishers, New York.

Zabalza, A. and Tzannatos, Z. (1985). The effect of Britain's anti-discriminatory legislation on relative pay and employment. *The Economic Journal*, 95(379):679–699.

Appendices

3.A Definitions, Data Sources and Extended Results

3.A.1 Definition of Four Rights Indices

Table A.1 summarizes the rights underlying each of the four indices and their sources.

Table A.1: Women's Rights Index Definitions

Index	Variable Included
Economic Rights	Whether a woman can apply for a passport in the same way as a man.[a]
	Whether a woman can travel outside the country in the same way as a man.[a]
	Whether a woman can travel outside her home in the same way as a man.[a]
	Whether a woman can choose where to live in the same way as a man.[a]
	Whether there is no legal provision that requires a married woman to obey her husband.[a]
	Whether a woman can be head of household in the same way as a man.[a]
	Whether a woman can obtain a judgment of divorce in the same way as a man.[a]
	Whether a woman has the same rights to remarry as a man.[a]
	Whether a woman can sign a contract in the same way as a man.[a]
	Whether a woman can register a business in the same way as a man.[a]
	Whether a woman can open a bank account in the same way as a man.[a]
	Whether the law prohibits discrimination in access to credit based on gender.[a]
	Whether men and women have equal ownership rights to immovable property.[a]
	Whether sons and daughters have equal rights to inherit assets from their parents.[a]
	Whether male and female surviving spouses have equal rights to inherit assets.[a]
	Whether the law grants male and female spouses equal administrative authority over assets during marriage.[a]
Political Rights	Whether there is female suffrage.[b]
	Whether women have the right to stand for elections.[c]
	Whether there have ever been more than 20% of women in parliament.[c,d]
Labor Rights	Whether a woman can get a job in the same way as a man.[a]
	Whether the law prohibits discrimination in employment based on gender.[a]
	Whether the law mandates equal remuneration for work of equal value.[a]
	Whether a woman can work at night in the same way as a man.[a]
	Whether a woman can work in a job deemed dangerous in the same way as a man.[a]
	Whether a woman can work in an industrial job in the same way as a man.[a]
Body Rights	Whether there is legislation on sexual harassment in employment.[a]
	Whether there are criminal penalties or civil remedies for sexual harassment in employment.[a]
	Whether there is legislation specifically addressing domestic violence.[a]
	Whether abortion is legal for physical health reasons[e,f,g]
	Whether abortion is legal for reasons other than physical health[e,f,g]
	Whether contraceptive prevalence among married women is equal to or above the cross country and time mean of $\approx 45\%$[h]

Data sources: [a]Women, Business and the Law database World Bank (2022a); [b]Skaaning et al. (2015); [c]Paxton et al. (2008); [d]World Bank (2022b); [e]based on Boyle et al. (2015); [f] United Nations (2022b); [g] United Nations (2022c); [h]United Nations (2022a)

3.A.2   Data Sources

Table A.2: Measurement and Sources of Independent Variables

Variable	Description	Source
GDP p.c. (in 1000s)	GDP per capita (constant 2015 USD)	World Bank (2022b)
Total fertility rate	Fertility rate, total (births per woman)	World Bank (2022b)
FLFP (15-64)	Labor force participation rate, female (% of female population ages 15-64)	World Bank (2022b), ILO (2022)
Majority catholic/muslim/buddhist	Dummy variable. 1 = more than 50% of the population is catholic/muslim/buddhist	Cline Center for Democracy (2022)
Secondary school enrollment	School enrollment, secondary (% gross)	World Bank (2022b)
Female population share	Population, female (% of total population) (constructed from male population share)	World Bank (2022b)
Employment in agriculture	Employment in agriculture (% of total employment) (modeled ILO estimate)	World Bank (2022b)
OPEC indicator	Dummy variable. 1 = Country is OPEC member	OPEC (2022)
Government effectiveness	QoG government effectiveness index combining the quality of public service provision, the quality of the bureaucracy, the competence of civil servants, the independence of the civil service from political pressures, and the credibility of the government's commitment to policies.	Ebetürk (2021), based on Teorell et al. (2013)
International pressure	Average number of ratifications related to women and children in the region (8 world regions)	Ebetürk (2021)
Nb. of conventions ratified	Total number of ratifications of treaties related to women's/ children's rights	Ebetürk (2021)
Women INGOs	Number of women INGOs in which residents of a country are members	Ebetürk (2021)
Women in parliament >20%	Dummy variable. 1 = Share of women in parliament has ever been larger than 20%	Paxton et al. (2008), World Bank (2022b)
Independence	Dummy variable. 1 = Country became independent in this year	Based on Boyle et al. (2015)

3.A.3   Changing Countries over the Sample Period

Between 1970 and 2019, some federal unions dissolved and new countries emerged. In other cases, countries merged. We briefly summarize how the WBL database treats such cases. Note that the database generally assumes that a hypothetical woman in question resides in the main business city of the country in question.

- If a country gained independence during the sample period, it was assigned the score of the country of which it was part prior to independence.

- For federal unions that dissolved over the sample period (for example, Yugoslavia or the USSR), the federal law was applied. The exception being if there was a law at the national

level of the constituent republic. In the WBL database, there are several instances where this is the case, even before independence.

- For formerly colonized countries, such as Antigua and Barbuda and Namibia, national laws generally applied during colonization.

- For Germany pre-1990, the legal score is based on the applicable law of Berlin, West Germany.

3.A.4   Regression Results

Table A.3 to Table A.6 show the extended regression results, discussed in Section 3.4.2.

Table A.3: Economic Rights

	Dependent Variable: Economic Rights Index					
	(1)	(2)	(3)	(4)	(5)	(6)
A. Economic Variables						
GDP p.c. (in 1000s)	0.32***	−0.16**	0.04	−0.17**	−0.05	
	(0.10)	(0.07)	(0.03)	(0.07)	(0.05)	
Total Fertility Rate		−7.15***	−1.71***	−6.75***	−6.53***	
		(0.56)	(0.28)	(0.63)	(0.62)	
Female LFP (15-64)		0.53***	0.14***	0.45***	0.43***	
		(0.06)	(0.03)	(0.07)	(0.07)	
Majority Catholic				0.74	−0.20	0.13
				(2.15)	(3.07)	(2.97)
Majority Muslim				−7.73**	−2.82	−22.24***
				(3.57)	(3.53)	(3.41)
Majority Buddhist				0.87	3.54	7.11**
				(4.05)	(5.04)	(3.26)
B. Interactions with Religion						
Majority Catholic × GDP p.c.					0.12	
					(0.11)	
Majority Muslim × GDP p.c.					−0.75***	
					(0.13)	
Majority Buddhist × GDP p.c.					−0.26	
					(0.20)	
Time Fixed Effects	No	No	Yes	Yes	Yes	Yes
Country Fixed Effects	No	No	Yes	No	No	No
Observations	7,796	5,428	5,428	4,045	4,045	6,709
Adjusted R^2	0.057	0.513	0.919	0.541	0.583	0.286

Notes: All regressions include a dummy variable that is equal to one for the year in which a country newly gained independence. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Column (1) uses GDP per capita as the only explanatory variable. Column (2) adds the total fertility rate and female labor force participation as additional economic variables potentially associated with women's rights. Column (3) introduces time and country fixed effects. Column (4) replaces country fixed effects with religion majority indicators to proxy cultural differences across countries. Column (5) interacts the religion majority indicators with GDP per capita. Column (6) controls for religion and time effects only.

Table A.4: Political Rights

	Dependent Variable: Political Rights Index					
	(1)	(2)	(3)	(4)	(5)	(6)
A. Economic Variables						
GDP p.c. (in 1000s)	0.11	−0.06	−0.06	−0.07	0.17*	
	(0.12)	(0.11)	(0.06)	(0.13)	(0.09)	
Total Fertility Rate		−2.45***	−1.36***	−1.75**	−1.51**	
		(0.64)	(0.44)	(0.72)	(0.63)	
Female LFP (15-64)		0.34***	0.35***	0.30***	0.26***	
		(0.07)	(0.04)	(0.09)	(0.08)	
Majority Catholic				1.95	5.11*	1.35
				(2.36)	(2.62)	(2.34)
Majority Muslim				−0.89	7.68**	−9.64**
				(3.79)	(3.32)	(3.98)
Majority Buddhist				−6.54	−1.90	−3.65
				(4.13)	(5.06)	(4.09)
B. Interactions with Religion						
Majority Catholic × GDP p.c.					−0.22	
					(0.16)	
Majority Muslim × GDP p.c.					−1.24***	
					(0.21)	
Majority Buddhist × GDP p.c.					−0.42**	
					(0.21)	
Time Fixed Effects	No	No	Yes	Yes	Yes	Yes
Country Fixed Effects	No	No	Yes	No	No	No
Observations	7,138	5,167	5,167	3,981	3,981	6,452
Adjusted R^2	0.009	0.140	0.757	0.181	0.305	0.143

Notes: All regressions include a dummy variable that is equal to one for the year in which a country newly gained independence. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Column (1) uses GDP per capita as the only explanatory variable. Column (2) adds the total fertility rate and female labor force participation as additional economic variables potentially associated with women's rights. Column (3) introduces time and country fixed effects. Column (4) replaces country fixed effects with religion majority indicators to proxy cultural differences across countries. Column (5) interacts the religion majority indicators with GDP per capita. Column (6) controls for religion and time effects only.

Table A.5: Labor Rights

	Dependent Variable: Labor Rights Index					
	(1)	(2)	(3)	(4)	(5)	(6)
A. Economic Variables						
GDP p.c. (in 1000s)	0.62***	0.26**	−0.26***	0.35***	0.60***	
	(0.14)	(0.10)	(0.06)	(0.11)	(0.10)	
Total Fertility Rate		−4.78***	2.24***	−2.59***	−2.37***	
		(0.86)	(0.48)	(0.93)	(0.90)	
Female LFP (15-64)		0.52***	0.06	0.31***	0.28***	
		(0.08)	(0.05)	(0.09)	(0.09)	
Majority Catholic				6.93*	10.63**	1.72
				(3.82)	(5.09)	(3.87)
Majority Muslim				−13.82***	−5.07	−22.97***
				(4.82)	(4.53)	(4.10)
Majority Buddhist				−5.66	0.61	−5.67
				(4.31)	(4.84)	(5.14)
B. Interactions with Religion						
Majority Catholic × GDP p.c.					−0.27	
					(0.31)	
Majority Muslim × GDP p.c.					−1.23***	
					(0.22)	
Majority Buddhist × GDP p.c.					−0.69***	
					(0.19)	
Time Fixed Effects	No	No	Yes	Yes	Yes	Yes
Country Fixed Effects	No	No	Yes	No	No	No
Observations	7,796	5,428	5,428	4,045	4,045	6,709
Adjusted R^2	0.123	0.299	0.826	0.380	0.441	0.264

Notes: All regressions include a dummy variable that is equal to one for the year in which a country newly gained independence. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Column (1) uses GDP per capita as the only explanatory variable. Column (2) adds the total fertility rate and female labor force participation as additional economic variables potentially associated with women's rights. Column (3) introduces time and country fixed effects. Column (4) replaces country fixed effects with religion majority indicators to proxy cultural differences across countries. Column (5) interacts the religion majority indicators with GDP per capita. Column (6) controls for religion and time effects only.

Table A.6: Body Rights

	Dependent Variable: Body Rights Index					
	(1)	(2)	(3)	(4)	(5)	(6)
A. Economic Variables						
GDP p.c. (in 1000s)	0.57***	0.14*	0.27***	0.22**	0.33***	
	(0.13)	(0.08)	(0.07)	(0.09)	(0.11)	
Total Fertility Rate		−5.82***	2.66***	−4.30***	−4.06***	
		(0.60)	(0.62)	(0.65)	(0.65)	
Female LFP (15-64)		0.32***	0.10**	0.20***	0.18**	
		(0.06)	(0.05)	(0.07)	(0.07)	
Majority Catholic				2.68	2.13	2.25
				(2.77)	(3.63)	(3.09)
Majority Muslim				−1.25	1.99	−8.22***
				(3.57)	(3.77)	(2.93)
Majority Buddhist				−3.70	−1.72	0.51
				(4.40)	(5.73)	(4.25)
B. Interactions with Religion						
Majority Catholic × GDP p.c.					0.07	
					(0.17)	
Majority Muslim × GDP p.c.					−0.52***	
					(0.12)	
Majority Buddhist × GDP p.c.					−0.18	
					(0.24)	
Time Fixed Effects	No	No	Yes	Yes	Yes	Yes
Country Fixed Effects	No	No	Yes	No	No	No
Observations	5,578	3,950	3,950	3,492	3,492	5,530
Adjusted R^2	0.106	0.268	0.776	0.430	0.445	0.335

Notes: All regressions include a dummy variable that is equal to one for the year in which a country newly gained independence. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Column (1) uses GDP per capita as the only explanatory variable. Column (2) adds the total fertility rate and female labor force participation as additional economic variables potentially associated with women's rights. Column (3) introduces time and country fixed effects. Column (4) replaces country fixed effects with religion majority indicators to proxy cultural differences across countries. Column (5) interacts the religion majority indicators with GDP per capita. Column (6) controls for religion and time effects only.

3.A.5 Graphical Illustration of Time Fixed Effects

Figure A.1: Time Fixed Effects by Rights Index
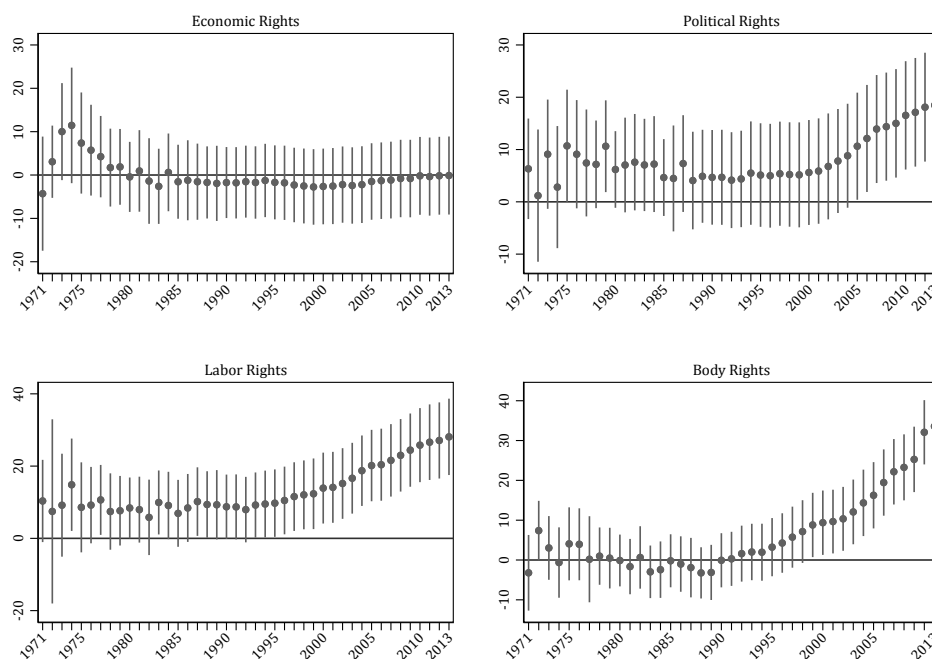
Notes: We plot the time fixed effects from Column (5) of Tables 3.3 to 3.6.

Supplementary data

Supplementary data are available at *JEEA* online.

Bibliography

Boyle, E. H., Kim, M., and Longhofer, W. (2015). Abortion liberalization in world society, 1960–2009 – dataset. *provided by the authors*.

Cline Center for Democracy (2022). Composition of religious and ethnic groups (CREG) project. https://clinecenter.illinois.edu/project/Religious-Ethnic-Identity/composition-religious-and-ethnic-groups-creg-project (accessed July 12, 2022).

Ebetürk, I. (2021). Global diffusion of laws: The case of minimum age of marriage legislation, 1965–2015 – dataset. *provided by the author*.

ILO (2022). Ilostat data explorer. https://www.ilo.org/shinyapps/bulkexplorer30/ (accessed August 9, 2022).

OPEC (2022). Member countries. https://www.opec.org/opec_web/en/about_us/25.htm (accessed September 14, 2022).

Paxton, P., Green, J., and Hughes, M. M. (2008). Women in parliament, 1945-2003: Cross-national dataset. https://doi.org/10.3886/ICPSR24340.v1. Inter-university Consortium for Political and Social Research.

Skaaning, S.-E., Gerring, J., and Bartusevicius, H. (2015). A lexical index of electoral democracy. *Comparative Political Studies*, 48(12):1491–1525.

Teorell, J., Sundström, A., Holmberg, S., Rothstein, B., Alvarado Pachon, N., and Dalli, C. M. (2013). The quality of government basic dataset. University of Gothenburg: The Quality of Government Institute. https://www.gu.se/en/quality-government.

United Nations (2022a). Model-based estimates and projections of family planning indicators 2022. Department of Economic and Social Affairs, Population Division. https://population.un.org/dataportal/home (accessed November 12, 2022).

United Nations (2022b). World population policies database (2011, 2013, 2015). https://esa.un.org/PopPolicy/wpp_datasets.aspx (accessed August 9, 2022).

United Nations (2022c). World population policies database (2017). https://www.un.org/development/desa/pd/data/world-population-policies (accessed August 9, 2022).

World Bank (2022a). Women business and the law data for 1971–2022. https://wbl.worldbank.org/content/dam/sites/wbl/documents/2021/02/WBL1971-2022%20Dataset.dta (accessed June 7, 2022).

World Bank (2022b). World development indicators. https://databank.worldbank.org/source/world-development-indicators (accessed July–August 2022).

Declaration

This dissertation is the result of my own work, and no other sources or means, except the ones listed, have been employed.

Laura Montenbruck

Curriculum Vitae

2018 - 2024	University of Mannheim
	Ph.D. in Economics
2016 - 2018	University of Göttingen
	M.A. in Development Economics
2011 - 2016	University of Passau
	B.A. in International Cultural and Business Studies
	B.Sc. in Business Administration and Economics